ISSN 2225-5435

Вестник УрФО.

БЕЗОПАСНОСТЬ В ИНФОРМАЦИОННОЙ СФЕРЕ

№ 2(60) / 2026

Подписной индекс 73852 в каталоге «Почта России»

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций.

Регистрационный номер ПИ № ФС77-91645 от 27.05.2026

Адрес редакции и издателя: Россия, 454080, г. Челябинск, пр. Ленина, д. 76. ЮУрГУ, Издательский центр
Тел./факс (351) 267-97-01.

Электронная версия журнала в Интернете:
www.info-secur.ru,
e-mail: urvest@mail.ru

16+

Journal of the Ural Federal District. Information security

ISSN 2225-5435

№ 2(60) / 2026

FOUNDER

SOUTH URAL STATE UNIVERSITY (NIU)

CHAIRMAN OF THE EDITORIAL BOARD

CHUVARDIN O. P.,
Head of Department Federal Service for Technical and Export Control of Russia for the Urals Federal District

CHIEF EDITOR

SOKOLOV A.N.,
Ph.D., Associate Professor, Head of Department "Information Protection", South Ural State University (National Research University) (Chelyabinsk city)

PRODUCING EDITOR

SOGRIN E. K.

LAYOUT

SCHREIBER A. E.

PROOFREADING

FEDOROV V. S.

Subscription index 73852

in the «Russian Post» catalog

The journal is registered by the Federal service in the field of communication, information technology and mass communications.

Registration number
PI No. ФС77-91645 dd. 27.05.2026

Editorial and publisher address: Russia, 454080, Chelyabinsk, Lenin Avenue, 76 SUSU, Publishing Center
Phone / fax (351) 267-97-01.

Electronic version of the magazine in the Internet:

www.info-secur.ru,
e-mail: urvest@mail.ru

16+

В НОМЕРЕ

РАДИОТЕХНИКА, В ТОМ ЧИСЛЕ СИСТЕМЫ И УСТРОЙСТВА ТЕЛЕВИДЕНИЯ

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ, ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

RADIO ENGINEERING, INCLUDING TELEVISION SYSTEMS AND DEVICES

METHODS AND SYSTEMS OF INFORMATION PROTECTION, INFORMATION SECURITY

УДК 004.056: 621.3

Вестник УрФО № 2(60) / 2026, с. 5–13

Хорев А.А., Чачилло Т.В

DOI: 10.14529/secur260201

ОЦЕНКА ВОЗМОЖНОСТИ ПЕРЕХВАТА РЕЧЕВОЙ ИНФОРМАЦИИ МЕТОДОМ «ВЫСОКОЧАСТОТНОЙ НАКАЧКИ»

В статье приведены результаты экспериментальных исследований по оценке возможностей перехвата речевой информации методом «высокочастотной накачки» телефонного аппарата при измерении уровней информативных сигналов на выходе полосового фильтра и на выходе аналогового амплитудного демодулятора анализатора спектра. Исследования проводились для RBW: 30 Гц и 10 Гц и VBW: 30 кГц, 10 кГц и 3 кГц. При проведении экспериментальных исследований установлено, что словесная разборчивость речи на выходе аналогового демодулятора анализатора спектра при расположении антенны на удалении 1 м от телефонного аппарата не превышает 26,7% даже для очень громкой речи (L = 84 дБ). Отношение сигнал/шум для информативных составляющих на выходе полосового фильтра анализатора достаточное для обеспечения высокой разборчивости речи при использовании цифровых методов демодуляции сигналов. При этом, изменение полосы пропускания RBW не оказывает существенного влияния на разборчивость речи. В реальных условиях перехват речевой информации из помещения методом «высокочастотной накачки» телефонного аппарата возможен только в случае его установки вблизи звуковых колонок систем звукоусиления помещений. При этом словесная разборчивость более 30% может обеспечиваться на расстояниях до 16 – 17 м.

Ключевые слова: *технический канал утечки акустической речевой информации, «высокочастотная накачка», разборчивость речи.*

ASSESSMENT OF THE POSSIBILITY OF INTERCEPTING SPEECH INFORMATION USING THE «HIGH-FREQUENCY PUMPING» METHOD

The article presents the results of experimental studies to evaluate the possibilities of intercepting speech information using the "high-frequency pumping" method of a telephone device when measuring the levels of informative signals at the output of a bandpass filter and at the output of an analog amplitude demodulator of a spectrum analyzer. The studies were conducted for RBW: 30 Hz and 10 Hz and VBW: 30 kHz, 10 kHz and 3 kHz. When conducting experimental studies, it was found that the verbal intelligibility of speech at the output of the analog demodulator of the spectrum analyzer when the antenna is located at a distance of 1 m from the telephone does not exceed 26.7% even for very loud speech (L = 84 dB). The signal-to-noise ratio for the informative components at the output of the bandpass filter of the analyzer is sufficient to ensure high speech intelligibility when using digital signal demodulation methods. At the same time, changing the RBW bandwidth does not significantly affect speech intelligibility. In real conditions, the interception of speech information from a room using the "high-frequency pumping" method of a telephone is possible only if it is installed near the sound speakers of indoor sound reinforcement systems. At the same time, verbal intelligibility of more than 30% can be achieved at distances up to 16-17 m.

Keywords: *technical channel of acoustic speech information leakage, "high-frequency pumping", speech intelligibility.*

Для перехвата акустической речевой информации из выделенных помещений (ВП) широко используются активные методы перехвата информации, такие как «высокочастотное навязывание», «высокочастотное облучение» и «высокочастотная накачка» («высокочастотная прокачка»).

Технические каналы утечки информации, реализуемые с использованием методов «высокочастотного навязывания» и «высокочастотного облучения», довольно подробно рассмотрены в работах [1–8], а методы и методики контроля подверженности технических средств акустоэлектрическим и акустоэлектромагнитным преобразованиям – в работах [9 – 15].

Технический канал утечки информации, создаваемый методом «высокочастотной накачки» технических средств (ТС), установленных в выделенном помещении (ВП), в доступной литературе впервые был описан в [16].

В работах [17, 18] приведены методика контроля подверженности технических средств «высокочастотной накачки» и результаты экспериментальных исследований подверженности телефонного аппарата «VEF TA-D» (TA611D) «высокочастотной накачке».

В результате исследований [17, 18] установлено, что телефон подвержен ВЧ-накачке в диапазоне частот от 30 МГц до 270 МГц, а максимальный уровень напряженности поля информативного сигнала наблюдается на частоте 230 МГц.

Для измерения характеристик переизлученного телефонным аппаратом информативного высокочастотного сигнала использовался анализатор FSH8. Измерения уровней информативных сигналов проводились на выходе полосового фильтра при отключенном предусилителе и полосе пропускания RBW = 30 Гц.

Цели данных экспериментальных исследований:

а) оценить разборчивость речи при ее

перехвате методом «высокочастотной накачки» телефонного аппарата при измерении уровней информативных сигналов на выходе полосового фильтра и на выходе аналогового амплитудного демодулятора анализатора спектра при различных полосах пропускания RBW и VBW;

б) разработать методический подход к оценке возможностей перехвата речевой информации методом «высокочастотной накачки» с учетом шумов антенны и приемного устройства, а также шумов телефонного аппарата.

Для проведения экспериментальных исследований был создан лабораторный стенд в составе: телефонный аппарат «VEF TA-D»; генератор сигналов «Rohde&Schwarz SMB100A»; электрическая дипольная активная измерительная антенна П6-51; анализатор спектра «Rohde&Schwarz FSL3» с опциями R&S®FSL-B7, R&S®FSL-B22 и R&S®FSL-K7; акустическая система «BEHRINGER B208D» с генератором НЧ-сигналов «Agilent 33521A»; шумомер «Эко физика-110А»; селективный микровольтметр «В6-17».

Учитывая, что 1 и 7 октавные полосы практически не влияют на разборчивость речи [1, 19] экспериментальные исследования проводились в 2 – 6 октавных полосах. Исследования проводились в соответствии с методикой, изложенной в [17]. В качестве тестового сигнала использовались гармонические колебания на среднегеометрических частотах пяти октавных полос: 250, 500, 1000, 2000 и 4000 Гц.

Измерения уровней информативных сигналов проводились на частоте высокочастотного сигнала 230 МГц при максимальной мощности высокочастотного сигнала (25 дБм) и максимальной громкости тестового акустического сигнала.

Исследования проводились в три этапа.

На **первом этапе** проводились исследования влияния полосы пропускания RBW анализатора спектра на разборчивость речи.

Измерения уровней информативных сигналов проводились при значениях полосы пропускания RBW: 30 Гц и 10 Гц.

В качестве тестовой системы использовалась акустическая система «BEHRINGER B208D» с генератором низкочастотных сигналов «Agilent 33521A».

Пример спектрограмм переизлученного телефонным аппаратом сигналов приведен на рис. 1.

Далее на основе результатов измерений уровней сигналов и шумов проводился расчет уровня информативного сигнала и отношения сигнал/шум в октавных полосах по формулам:

$$U_{c.i} = 10\,lg(\,10^{0,1U_{(c+ш),i}} - 10^{0,1U_{ш.г.i}}); \qquad (1)$$

$$q_{c.i} = 20\lg\left[\frac{10^{0,05\cdot U_{c.i}\cdot\sqrt{RBW}}}{10^{0,05\cdot U_{ш.г.i}\cdot\sqrt{\Delta F_i}}}\right], \qquad (2)$$

где $U_{c.i}$ – уровень информативного сигнала на среднегеометрической частоте i-й октавной полосы, дБ(мкВ);

$U_{(с+ш).i}$ – уровень модуляционной составляющей при включенном тестовом сигнале на среднегеометрической частоте i-й октавной полосы, дБ(мкВ);

$U_{ш.г.i}$ – уровень шумов при включенном тестовом сигнале на среднегеометрической частоте i-й октавной полосы и включенном высокочастотном генераторе, дБ(мкВ);

$q_{c.i}$ – отношение сигнал/шум в i-й октавной полосы, дБ;

ΔF_i – ширина i-й октавной полосы, Гц;

RBW – ширина полосы пропускания на выходе полосового фильтра анализатора спектра, Гц.

Рассчитанные по формуле (2) отношения сигнал/шум в октавных полосах для максимальных уровней тестового сигнала приведены в табл.1.

Отношения сигнал/шум, приведенные в табл. 1, рассчитаны для уровней тестового акустического сигнала, существенно превышающих уровни типовых речевых сигналов [1]. Проведенные исследования [17, 18] показали, что при перехвате информации методом «высокочастотной накачки», отношение сигнал/шум на входе приемного устройства растет прямо пропорционально уровню тестового акустического сигнала. Следовательно, для расчета отношения сигнал/шум, соответствующего типовым значениям уровней сигналов в октавных полосах, можно рассчитать по формуле:

$$q_{c.m.i} = q_{c.i} - \Delta L = q_{c.i} - (L_{и.i} - L_{c.i}), \qquad (3)$$

где $q_{c.i}$ – отношение сигнал/шум для i-й октавной полосы, соответствующее уровню тестового акустического сигнала, дБ;

$q_{c.m.i}$ – отношение сигнал/шум для i-й октавной полосы, соответствующее типовому спектру речевого сигнала, дБ;

$L_{и.i}$ – измеренный уровень громкости тестового сигнала в i -й октавной полосе, дБ;

$L_{c.i}$ – типовой уровень громкости речевого сигнала в i -й октавной полосе [1], дБ.

Результаты расчетов отношений сигнал/

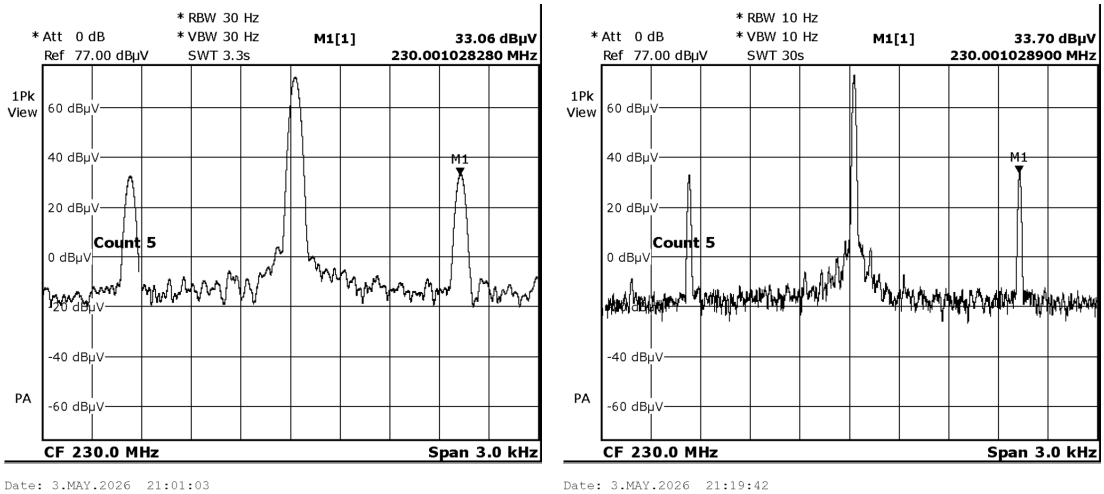

Рис.1. Спектрограммы переизлученных телефонным аппаратом модулированных высокочастотных сигналов (частота сигнала ВЧ-навязывания – 230 МГц, мощность генератора ВЧ-сигнала – 25 дБм, уровень тестового акустического сигнала – максимальный, частота тестового акустического сигнала – 1 кГц): RBW= 30 Гц (а), RBW=10 Гц (б)

Таблица 1

Рассчитанные отношения сигнал/шум в октавных полосах

Номер октавной полосы (диапазон частот, Гц)	Уровень тестового акустического сигнала ($L_{и.i}$), дБ	Отношение сигнал/шум ($q_{c.i}$), дБ	
		RBW = 30 Гц	RBW = 10 Гц
2 (180 – 355)	109,8	19,3	17,2
3 (355 – 710)	110,5	24,5	29,1
4 (710 – 1420)	105,4	33,1	32,9
5 (1420 – 2840)	102,7	21,2	25,9
6 (2840 –5680)	105,7	31,6	27,7

Таблица 2

Рассчитанные отношения сигнал/шум в октавных полосах, приведенные к типовым уровням громкости, и результаты расчетов словесной разборчивости речи (W)

№ октавной полосы (диапазон частот, Гц)	Отношение сигнал/шум ($q_{c.m.i}$), дБ, при типовом уровне речевого сигнала (L), дБ					
	RBW = 30 Гц			RBW = 10 Гц		
	L = 70	L = 76	L = 84	L = 70	L =76	L = 84
2 (180 – 355)	-24,5	-18,5	-10,5	-26,6	-20,6	-12,6
3 (355 – 710)	-20,0	-14,0	-6,0	-15,4	-9,38	-1,38
4 (710 – 1420)	-11,3	-5,3	2,7	-11,5	-5,5	2,5
5 (1420 – 2840)	-25,5	-19,5	-11,5	-20,9	-14,9	-6,9
6 (2840 –5680)	-21,1	-15,1	-7,1	-25,0	-19,0	-11,0
W, %	8,3	27,1	71,4	8,5	28,5	74,4

шум и словесной разборчивости речи, проведенной в соответствии с методикой, изложенной в [1], для речи средней громкости (L_c = 70 дБ), громкой речи (L_c = 76 дБ) и очень громкой речи (L_c = 84 дБ) приведены в табл. 2.

Анализ данных, представленных в табл. 2, показал, что изменение полосы пропускания RBW не оказывает существенного влияния на разборчивость речи, но при полосе пропускания RBW = 10 Гц разборчивость речи незначительно выше, чем при RBW = 30 Гц.

На **втором этапе** проводились исследо-

вания влияния полосы пропускания VBW анализатора спектра на разборчивость речи.

Исследования также проводились в соответствии с методикой, изложенной в [17]. Однако, измерения отношений сигнал/шум проводились не на выходе полосового фильтра, а на выходе аналогового демодулятора с использованием селективного микровольтметра В6-17 при его полосе пропускания ΔF = 10 Гц.

Измерения уровней информативных сигналов проводились при трех значениях полосы пропускания VBW: 30 кГц, 10 кГц и 3 кГц.

Рассчитанные по формуле (2) отношения сигнал/шум в октавных полосах на аналоговом выходе демодулятора для максимальных уровней тестового сигнала приведены в табл. 3.

Рассчитанные отношения сигнал/шум в октавных полосах, приведенные к типовым

Таблица 3

Рассчитанные отношения сигнал/шум в октавных полосах

Номер октавной полосы (диапазон частот, Гц)	Уровень тестового акустического сигнала ($L_{и.i}$), дБ	Отношение сигнал/шум ($q_{г.i}$), дБ		
		VBW = 30 кГц	VBW = 10 кГц	VBW = 3 кГц
2 (180 – 355)	109,8	7,91	9,42	8,61
3 (355 – 710)	110,5	15,79	16,60	20,50
4 (710 – 1420)	105,4	11,11	16,11	10,01
5 (1420 – 2840)	102,7	23,64	20,14	18,14
6 (2840 –5680)	105,7	20,63	20,13	-27,37

Таблица 4

Рассчитанные отношения сигнал/шум в октавных полосах, приведенные к типовым уровням громкости, и результаты расчетов словесной разборчивости речи (W)

Номер октавной полосы (диапазон частот, Гц)	Отношение сигнал/шум ($q_{c.m.i}$), дБ, при типовом уровне речевого сигнала (L), дБ								
	VBW = 30 кГц			VBW = 10 кГц			VBW = 3 кГц		
	L=70	L=76	L = 84	L=70	L=76	L=84	L=70	L=76	L=84
2 (180 – 355)	-35,9	-29,9	-21,9	-34,4	-28,4	-20,4	-35,2	-29,2	-21,2
3 (355 – 710)	-28,7	-22,7	-14,7	-27,9	-21,9	-13,9	-24,0	-18,0	-10,0
4 (710 – 1420)	-33,3	-27,3	-19,3	-28,3	-22,3	-14,3	-34,4	-28,4	-20,4
5 (1420 – 2840)	-23,1	-17,1	-9,1	-26,6	-20,6	-12,6	-28,6	-22,6	-14,6
6 (2840 –5680)	-32,1	-26,1	-18,1	-32,6	-26,6	-18,6	-80,1	-74,1	-66,1
W, %	**1,0**	**5,2**	**26,7**	**0,4**	**2,8**	**19,9**	**0,2**	**1,3**	**10,2**

уровням громкости, и результаты расчетов словесной разборчивости речи, проведенной в соответствии с методикой, изложенной в [1], представлены в табл. 4.

Анализ данных, представленных в табл. 4, показал, что:

– разборчивость речи на выходе аналогового демодулятора существенно ниже, чем на выходе полосового фильтра. Очевидно, это обусловлено очень малой глубиной модуляции сигнала (от 0,2% – для 2-й октавной полосы и до 2,1% - для 6-й октавной полосы) и высоким уровнем шумов низкочастотного тракта;

– максимальная разборчивость речи наблюдается при VBW = 30 кГц.

Таким образом, в ходе экспериментальных исследований установлено, что наибольшая разборчивость речи наблюдается при измерении уровней модуляционных составляющих высокочастотных сигналов, переизлученных техническим средством, на выходе полосового фильтра при RBW = 10 Гц. Следовательно, именно данный метод измерений целесообразно использовать при оценке возможностей технических средств разведки (ТСР) по перехвату речевой информации методом «высокочастотной накачки» технических средств, установленных в выделенном помещении.

На **третьем этапе** проводилась оценка возможностей перехвата речевой информации из выделенного помещения методом «высокочастотной накачки» телефонного аппарата.

Результаты расчета словесной разборчивости, приведенной в табл. 2 и 4, основываются на использовании в качестве приемного устройства анализатора спектра «Rohde& Schwarz FSL3» с измерительной антенной П6-51, при ее установке на расстоянии 1 м от телефонного аппарата.

На расстоянии r = 1 м от ТС (телефонного аппарата) шумы на входе приемного устройства, обусловленные шумами ТС, существенно выше как собственных шумов приемного устройства, так и шумов антенны. Следовательно, для r = 1 м можно полагать, что $U_{ш.i} \approx \approx U_{ш.тс.i}$($U_{ш.тс.i}$ измеряется при включенном высокочастотном генераторе).

С увеличением расстояния, вследствие затухания электромагнитных излучений (ЭМИ), как напряжение информативного сигнала на входе приемного устройства, так и напряжение шумов, обусловленных шумами ТС, будут падать, но при этом напряжение шумов антенны, пересчитанное на вход приемного устройства, и напряжение собственных шумов приемного устройства будут оставаться неизменными.

Учитывая, что шумы антенны П6-51, пересчитанные на вход приемного устройства, значительно выше собственных шумов анализатора спектра FSL - 3, напряжение шумов на входе приемного устройства для i-й октавной полосы на расстоянии r от ТС можно рассчитать по формуле:

$$U_{ш.i.r} \approx 20lg\sqrt{10^{0,1\cdot(U_{ш.тс.i}-V_r)} + 10^{0,1\cdot(U_{ш.a.i})}}, \quad (4)$$

где $U_{ш.i.r}$ – напряжение шумов на входе приемного устройства для i-й октавной полосы на расстоянии r от ТС, дБ;

$U_{ш.a.i}$ – напряжение шумов антенны, пересчитанное на вход приемного устройства, для i-й октавной полосы, дБ;

$U_{ш.тс.i}$ – напряжение шумов на входе приемного устройства, обусловленных шумами технического средств (ТС), для i-й октавной полосы, измеренное на расстоянии r от ТС, дБ.

V_r – затухание ЭМИ на частоте сигнала f на расстоянии r от ТС, дБ.

Учитывая, что измерение ЭМИ проводилось на расстоянии r =1 м, для расчета затухания ЭМИ на расстоянии r от ТС воспользуемся формулой [20]:

$$V_r \approx 20lg\left(\frac{r^3}{\sqrt{\frac{k^4r^4-k^2r^2+1}{k^4-k^2+1}}}\right), \quad (5)$$

где V_r – затухание ЭМИ на расстоянии r, дБ;

$k = \frac{2\pi}{\lambda} = \frac{2\pi \cdot f}{300}$ – волновое число;

λ – длина волны излучения, м;

f – частота излучения, МГц;

r – расстояние от источника излучения, м.

С учетом формул (4) и (5) отношение сигнал/шум на входе приемного устройства можно рассчитать по формуле:

$$q_{c.m.i.r} = U_{c.m.i} - V_r - U_{ш.i.r}. \quad (6)$$

где $q_{c.m.i}$ – отношение сигнал/шум для i-й октавной полосы, соответствующее уровню тестового акустического сигнала, рассчитанное для расстояния r = 1 м, дБ;

$q_{c.m.i.r}$ – отношение сигнал/шум для i-й октавной полосы, соответствующее уровню тестового акустического сигнала, рассчитанное для расстояния r, дБ.

На основе данных, представленных в табл. 2 для RBW = 10 Гц, был выполнен расчет разборчивости речи на различных расстояниях. Результаты расчетов приведены в табл. 5

Анализ данных, представленных в табл. 5, показал, что в реальных условиях перехват речевой информации из помещения методом «высокочастотной накачки» телефонного аппарата возможен только в случае его установки вблизи звуковых колонок систем звукоусиления помещений. При этом словесная разборчивость более 30% может обеспечиваться на расстояниях до 16 – 17 м.

Заключение

Таким образом, анализ результатов проведенных исследований показал, что при использовании в качестве приемного устройства средства перехвата информации анализатора спектра FSL-3 с антенной П6-51:

– словесная разборчивость речи на выходе аналогового демодулятора анализатора спектра при расположении антенны на удалении 1 м от телефонного аппарата не превышает 26,7% даже для очень громкой речи (L = = 84 дБ). Очевидно, это обусловлено очень малой глубиной модуляции сигнала (менее 2%) и высоким уровнем шумов низкочастотного тракта. Максимальная разборчивость речи наблюдается при VBW = 30 кГц;

– отношение сигнал/шум для информативных (боковых) составляющих на выходе полосового фильтра анализатора достаточное для обеспечения высокой разборчивости речи при использовании цифровых методов демодуляции сигналов;

– изменение полосы пропускания RBW не оказывает существенного влияния на разборчивость речи, но при полосе пропускания RBW = 10 Гц разборчивость речи незначительно выше, чем при RBW = 30 Гц;

Таблица 5

Результаты расчетов разборчивости речи на различных расстояниях

Расстояние от технического средства	Разборчивость речи W (%) при типовом уровне речевого сигнала L		
	L = 70 дБ	L = 76 дБ	L = 84 дБ
r = 1 м	8,30	28,00	74,12
r = 5 м	6,09	22,04	65,32
r = 10 м	3,37	13,79	49,77
r = 15 м	1,85	8,54	36,89
r = 20 м	1,07	5,47	27,51
r = 25 м	0,65	3,64	20,84
r = 30 м	0,42	2,52	16,07
r = 35 м	0,28	1,79	12,60
r = 40 м	0,19	1,31	10,04
r = 45 м	0,14	0,99	8,11
r = 50 м	0,10	0,76	6,64

– на расстоянии r = 1 м от ТС шумы на входе приемного устройства, обусловленные шумами ТС, существенно выше как собственных шумов приемного устройства, так и шумов антенны. С увеличением расстояния от ТС, вследствие затухания ЭМИ, на входе приемного устройства напряжение информативного сигнала и напряжение шумов, обусловленных шумами ТС, будут падать, но при этом напряжение шумов антенны, пересчитанное на вход приемного устройства, и напряжение собственных шумов приемного устройства будут оставаться неизменными;

– в реальных условиях перехват речевой информации из помещения методом «высокочастотной накачки» телефонного аппарата возможен только в случае его установки вблизи звуковых колонок систем звукоусиления помещений. При этом словесная разборчивость более 30% может обеспечиваться на расстояниях до 16 – 17 м.

Литература

1. Хорев А.А. Техническая защита информации: учеб. пособие: В 3-х т. Т. 1: Технические каналы утечки информации. -М.: НПЦ «Аналитика», 2008. – 436 с.

2. Лысов А.В. Высокочастотное зондирование акустических возбуждаемых объектов в проводной среде. Кабельные локационные системы акустической разведки. – СПб.: Медиапапир, 2021. – 628 с.

3. Лысов А.В. Электромагнитное зондирование акустических возбуждаемых объектов (радиолокационные системы акустической разведки). – СПб.: Медиапапир, 2020. – 678 с.

4. Авдеев В.Б. Способ дистанционного перехвата речевой информации из защищаемого помещения здания с охраняемой зоной. Патент на изобретение № 2 575 406. Дата регистрации 20.02.2016 г. Заявка № 2014145030/08 от 06.11.2014 г.

5. Авдеев В.Б., Анищенко А.В. Способ дистанционного перехвата конфиденциальной речевой информации, циркулирующей в защищенном помещении. Патент на изобретение № 2 642 034. Дата регистрации 23.01.2018 г. Заявка № 2016129825 от 20.07.2016 г.

6. Авдеев В.Б., Катруша А.Н. Способ дистанционного перехвата речевой информации из защищаемого помещения. Патент на изобретение № 2 558 673. Дата регистрации 10.08.2015 г. Заявка № 2014137732/07 от 17.09.2014 г.

7. Авдеев В.Б., Катруша А.Н. Способ радиоперехвата речевой информации из защищаемого помещения. Патент на изобретение № 2 561 507. Дата регистрации 27.08.2015 г. Заявка № 2014142671/07 от 22.10.2014 г.

8. Колесник Д. Ю., Майоров А. И. Высокочастотное навязывание. Принцип реализации и методы защиты// Технические средства защиты информации: тезисы докладов XV Белорусско-российской науч.-техн. конф. (Минск, 6 июня 2017 г.). – Минск: БГУИР, 2017. – С. 31 – 32.

9. Кондратьев А. В. Техническая защита информации. Практика работ по оценке основных каналов утечки: учеб. пособие. – М.: Горячая линия – Телеком, 2016. – 304 с.

10. Лыков Ю. В., Морозова А. Д., Кукуш В. Д., Парфёнов А. С. Исследование метода ВЧ-навязывания для несанкционированного съема информации с телефонных линий//ScienceRise. – Харьков: 2015. – том 7, №2 (12). – С. 51 – 56.

11. Хорев А.А., Лукманова О.Р. Математическое моделирование акустоэлектрического канала утечки речевой информации, создаваемого методом «высокочастотного навязывания»//Защита информации. Инсайд. – С. Петербург: 2023. – № 1 (109) – С. 3– 11.

12. Шестаков И.И. Помехоустойчивость выделения информационного сигнала из интермодуляционного излучения при высокочастотном навязывании//Вестник СибГУТИ. – Новосибирск: 2011. – № 3. – С. 45- 58.

13. Дураковский А.П., Куницын И.В., Лаврухин Ю.Н. Контроль защищенности речевой информации в помещениях. Аттестационные испытания вспомогательных технических средств и систем по требованиям безопасности информации: учеб. пособие. – М.: НИЯУ МИФИ, 2015. – 152 с.

14. Хорев А.А. Контроль защищенности вспомогательных технических средств и систем от утечки по акустоэлектрическим каналам// Специальная техника. – М.: 2014. – № 6 – С. 48 – 63

15. Хорев А.А., Лукманова О.Р. Контроль подверженности телефонного аппарата «высокочастотному навязыванию» с использованием виртуального лабораторного стенда//Международная конференция «Радиоэлектронные устройства и системы для инфотелекоммуникационных технологий – РЭУС-2018». Доклады. – М.: РНТОРЭС имени А.С. Попова. 2018. – С. 343 – 348.

16. Катруша А.Н. Обобщенный метод анализа параметрических каналов перехвата акустической речевой информации// «Воздушно-космические силы. Теория и практика». – Воронеж, 2024. – № 31. – С. 122 – 133. ISSN: 2500 – 4352.

17. Хорев А.А., Чачилло Т. В. Исследование подверженности телефонного аппарата «высокочастотной накачке» // Вестник УрФО. Безопасность в информационной сфере. – 2025. – № 2(56). – С. 80–95.

18. Хорев А.А., Чачилло Т.В. Исследование возможности перехвата акустической речевой информации методом «высокочастотной накачки»// Всероссийская конференция «Радиоэлектронные устройства и системы для инфотелекоммуникационных технологий – РЭУС-2025». Доклады. – М.: РНТОРЭС имени А.С. Попова. 2025. – С. 434 – 440. ISBN 978-5-905278-62-4.

19. Хорев А.А., Порсев И.С. Вероятностный метод оценки разборчивости речи // Всероссийская конференция «Радиоэлектронные устройства и системы для инфотелекоммуникационных технологий – РЭУС-2023». Доклады. – М.: РНТОРЭС имени А.С.Попова. 2023. – С. 422 – 427. ISBN 978-5-905278-54-9.

20. Авдеев В.Б., Катруша А.Н. Оценка уровня побочного электромагнитного излучения на границе контролируемой зоны с учетом продольной компоненты поля//Специальная техника. – М.: 2014. – № 1. С. 28–34. ISSN: 1996-0506.

References

1. Khorev A.A. Technical information protection: textbook. manual: In 3 volumes Vol. 1: Technical channels of information leakage. – M.: NPC "Analytics", 2008. – 436 p.

2. Lysov A.V. High-frequency sounding of acoustic excited objects in a wired environment. Acoustic reconnaissance cable location systems. – St. Petersburg: Mediapapir, 2021. – 628 p.

3. Lysov A.V. Electromagnetic sounding of acoustic excited objects (acoustic reconnaissance radar systems). – St. Petersburg: Mediapapir, 2020. – 678 p.

4. Avdeev V.B. A method of remote interception of speech information from a protected building with a protected area. Patent for invention No. 2,575,406. Registration date: 02/20/2016 Application no. 2014145030/08 dated 11/06/2014

5. Avdeev V.B., Anishchenko A.V. A method for remote interception of confidential speech information circulating in a secure room. Patent for invention No. 2,642,034. Registration date: 01/23/2018. Application No. 2016129825 dated 07/20/2016.

6. Avdeev V.B., Katrusha A.N. A method of remote interception of speech information from a protected room. Patent for invention No. 2,558,673. Registration date 08/10/2015 Application No. 2014137732/07 on 17.09.2014.

7. Avdeev V.B., Katrusha A.N. Method of radio interception of speech information from a protected room. Patent for invention No. 2,561,507. Registration date 08/27/2015 Application No. 2014142671/07 dated 22.10.2014

8. Kolesnik D. Yu., Mayorov A. I. High-frequency imposition. The principle of implementation and methods of protection// Technical means of information protection: abstracts of the XV Belarusian-Russian Scientific and Technical conference (Minsk, June 6, 2017). Minsk: BGUIR, 2017. pp. 31–32.

9. Kondratiev A.V. Technical protection of information. The practice of assessing the main leakage channels: textbook. the manual. – M.: Hotline - Telecom, 2016. – 304 p. 10. Lykov Yu. V., Morozova A.D., Kukush V. D., Parfenov A. S. Investigation of the HF-imposition method for unauthorized removal of information from telephone lines//ScienceRise. – Kharkiv: 2015. Volume 7, No. 2 (12). pp. 51–56.

11. Khorev A.A., Lukmanova O.R. Mathematical modeling of the acousto–electric channel of leakage of speech information created by the method of "high-frequency imposition"//Information protection. Inside. – St. Petersburg: 2023. – № 1 (109) – P– 3–11.

12. Shestakov I.I. Noise immunity of information signal isolation from intermodulation radiation during high-frequency imposition//Bulletin of SibGUTI. Novosibirsk: 2011. No. 3. pp. 45–58.

13. Durakovsky A.P., Kunitsyn I.V., Lavrukhin Yu.N. Control of the security of speech information in rooms. Certification tests of auxiliary technical means and systems for information security requirements: textbook. The manual. Moscow: NRU MEPhI, 2015. 152 p.

14. Khorev A.A. Monitoring the protection of auxiliary equipment and systems from leakage through acousto-electric channels// Special equipment. – M.: 2014. – № 6 – pp. 48–63

15. Khorev A.A., Lukmanova O.R. Control of the telephone's susceptibility to "high-frequency imposition" using a virtual laboratory stand//International Conference "Radio Electronic devices and systems for Infotelec communication Technologies – REUS-2018". Reports. Moscow: RNTORES named after A.S. Popov. 2018. – pp. 343–348.

16. Katrusha A.N. Generalized method of analysis of parametric channels of interception of acoustic speech information// "Aerospace forces. Theory and practice". Voronezh, 2024. No. 31. pp. 122–133. ISSN: 2500-4352.

17. Khorev A.A., Chachillo T. V. Investigation of the susceptibility of a telephone device to "high–frequency pumping" // Bulletin of the Ural Federal District. Information security. – 2025. – № 2(56). – Pp. 80–95.

18. Khorev A.A., Chachillo T.V. Investigation of the possibility of intercepting acoustic speech information using the "high–frequency pumping" method// All-Russian Conference "Radio electronic devices and systems for infotelec communication technologies – REUS-2025". Reports. Moscow: RNTORES named after A.S. Popov. 2025. – pp. 434-440. ISBN 978-5-905278-62-4

19. Khorev A.A., Porsev I.S. Probabilistic method of speech intelligibility assessment // All-Russian conference "Radioelectronic devices and systems for infotelec communication technologies – REUS-2023". Reports. Moscow: RNTORES named after A.S.Popov. 2023. pp. 422-427. ISBN 978-5-905278-54-9

20. Avdeev V.B., Katrusha A.N. Assessment of the level of side electromagnetic radiation at the boundary of the controlled zone, taking into account the longitudinal component of the field//Special equipment. – M.: 2014.– No. 1. pp. 28-34. ISSN: 1996–0506

ХОРЕВ Анатолий Анатольевич, доктор технических наук, профессор, заведующий кафедрой «Информационная безопасность» федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский университет «Московский институт электронной техники» (МИЭТ), 124498, г. Москва, г. Зеленоград, площадь Шокина, дом 1, МИЭТ, E-mail: horev@miee.ru

ЧАЧИЛЛО Тимофей Витальевич, студент кафедры «Информационная безопасность» федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский университет «Московский институт электронной техники» (МИЭТ). 124498, г. Москва, г. Зеленоград, площадь Шокина, дом 1, МИЭТ. E-mail: chachillo1502@gmail.com

KHOREV Anatoly Anatolyevich, Doctor of Technical Sciences, Professor, Head of the Department "Information Security" of the Federal State Autonomous Educational Institution of Higher Education "National Research University "Moscow Institute of Electronic Technology" (MIET). 124498, Moscow, Zelenograd, Shokin Square, house 1, MIET. E-mail: horev@miee.ru

CHACHILLO Timofey V., student of the Department of Information Security at the Federal State Autonomous Educational Institution of Higher Education National Research University Moscow Institute of Electronic Technology (MIET). Shokin Square, 1, MIET, Moscow, 124498, Zelenograd. E-mail: chachillo1502@gmail

УДК 004.934:316.6 Вестник УрФО № 2(60) / 2026, с. 14–23

Аракчаа А.М., Демин Д.М., Жусов Д.Л., Макеев С.М., Соколов А.Н.

DOI: 10.14529/secur260202

АНАЛИЗ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ МОДЕЛЕЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ РАСПОЗНАВАНИЯ РЕЧИ ПРИ ПРОВЕДЕНИИ СОЦИОЛОГИЧЕСКИХ ОПРОСОВ

Представлены результаты сравнительного анализа открытых моделей искусственного интеллекта для автоматического распознавания речи в условиях полевых социологических опросов. Целью исследования было определение наиболее эффективной модели семейства Whisper для транскрибации интервью при наличии акустических помех, характерных для уличных условий проведения опросов. Методология включает экспериментальную оценку шести версий модели Whisper (Tiny, Base, Small, Medium, Large-v3, Large-v3-turbo) с применением метрики IWER (Inflectional Word Error Rate). Моделирование проводилось на аудиофайлах с искусственно внесёнными шумами различной интенсивности (соотношение сигнал/шум от 9,9 до 22 LUFS), имитирующими пять типичных сценариев проведения опросов. Установлено, что модель Large-v3-turbo демонстрирует наилучшую точность распознавания: показатель метрики IWER составляет от 0,79% (условия жилого сектора) до 4,86% (непосредственно у дороги), что на 20% ниже по сравнению с базовой моделью Tiny. Новизна исследования заключается в использовании сравнительной оценки метрики IWER для моделей Whisper в условиях, специфичных для социологических полевых исследований. Полученные данные позволяют обосновать выбор модели для создания специализированных информационных систем автоматизированной обработки речевых ответов респондентов.

Ключевые слова: *социологический опрос, распознавание речи, искусственный интеллект, Whisper, метрика IWER, акустические помехи, транскрибация, полевые исследования, нейронные сети, автоматизация обработки данных.*

ANALYSIS OF THE POSSIBILITY OF USING ARTIFICIAL INTELLIGENCE MODELS FOR SPEECH RECOGNITION IN SOCIOLOGICAL SURVEYS

The results of a comparative analysis of open-source artificial intelligence models for automatic speech recognition under field sociological survey conditions are presented. The aim of the study was to determine the most effective model of the Whisper family for interview transcription in the presence of acoustic noise characteristic of outdoor survey conditions. The methodology includes an experimental evaluation of six versions of the Whisper model (Tiny, Base, Small, Medium, Large-v3, Large-v3-turbo) using the IWER (Inflectional Word Error Rate) metric. Modeling was performed on audio files with artificially added noise of varying intensity (signal-to-noise ratio ranging from 9.9 to 22 LUFS), simulating five typical survey scenarios. It was established that the Large-v3-turbo model demonstrates the highest recognition accuracy: the IWER metric ranges from 0.79% (residential area conditions) to 4.86% (directly near the road), which is 20% lower compared to the baseline Tiny model. The novelty of the research lies in the comparative evaluation of the IWER metric for Whisper models under conditions specific to sociological field research. The obtained data enable substantiating the model choice for creating specialized information systems for automated processing of respondents' speech responses.

Keywords: *sociological survey, speech recognition, artificial intelligence, Whisper, IWER metric, acoustic noise, transcription, field research, neural networks, data processing automation.*

В настоящее время известны различные области применения систем автоматического распознавания речи [1]: построение систем управления устройствами голосовыми командами, автоматизация обработки звонков и улучшения качества обслуживания клиентов, транскрибация аудиозаписей в текст, в т.ч. в условиях помех и др. Среди вышеназванных, отдельный интерес представляет обработка социологических данных, получаемых в ходе проведения социологических опросов в форме интервьюирования, где также могут применяться подобные системы [2].

Известны различные средства распознавания речи, среди которых можно выделить коммерческие (Google Speech-to-Text, Microsoft Azure Speech Service, Amazon Transcribe, IBM Watson Speech to Text и т.д.) [3, 4] и открытые (Whisper (OpenAI), Vosk, SpeechBrain и т.д.) [5]. Среди них наиболее точными для решения прикладных задач являются Google Speech-to-Text (работает на основе SpeechRecognition), Whisper, Vosk и SpeechBrain [5-8]. Сравнительная характеристика этих средств представлена в табл. 1.

Анализ данных табл. 1 свидетельствует о том, что наибольший научный интерес на текущем этапе представляют модели системы распознавания речи Whisper. К их ключевым преимуществам перед аналогами относятся поддержка русского языка, открытый доступ к архитектурам моделей и возможность автономной работы.

Особенности измерения и представления громкости звука в средствах вычислительной техники

При распознавании речи ключевой характеристикой в условиях четкой дикции интервьюера и респондента является громкость.

Таблица 1

Сравнительная характеристика средств распознавания

Параметры \ Средства распознавания	**Whisper**	**Speech Recognition**	**SpeechBrain**	**Vosk**
Режим работы	Офлайн/ Онлайн	Онлайн	Онлайн	Офлайн/Онлайн
Доступ к архитектурам моделей	Есть	Есть	Есть	Отсутствует
Распознавание русского языка	+	+	–	+

Таблица 2

Сравнительная таблица моделей Whisper

Модель	**Количество параметров**	**Область применения**
Tiny	39 млн	Быстрое распознавание речи с низкими требованиями к вычислительным ресурсам и точности
Base	74 млн	Задачи, требующие более качественного распознавания при ограниченных вычислительных ресурсах
Small	244 млн	Более сложные задачи распознавания речи, где требуется баланс между качеством и быстротой
Medium	769 млн	Профессиональные задачи, где важна высокая точность и качество распознавания с ограниченными вычислительными ресурсами
Large-v3	1,55 млрд	Крупномасштабные и профессиональные проекты, где важна максимальная точность и поддержка множества языков, при этом задействуются большие вычислительные ресурсы
Large-v3-turbo	809 млн	Задачи, где требуется быстрое и точное распознавание речи, с ограниченными вычислительными ресурсами

Громкость звука представляет собой субъективное восприятие силы звукового сигнала, которое коррелирует с интенсивностью звуковых колебаний и уровнем звукового давления. В окружающей среде громкость измеряется с помощью приборов, фиксирующих интенсивность звука непосредственно, например, при разговоре человека, когда его голос регистрируется специальными устройствами, и результат выражается в децибелах (дБ) [9]. В цифровых вычислительных системах, когда речь представлена в виде звукового файла, прямое измерение громкости в дБ невозможно, поскольку вычислительная система не осуществляет прямое измерение акустического сигнала. В таких условиях для оценки громкости применяется величина LUFS (Loudness Units relative to Full Scale), представляющая собой стандартизированный способ измерения средней громкости аудиосигнала с учетом чувствительности человеческого уха к различным частотам. Она предназначена для выравнивания уровня громкости звукового файла при воспроизведении на различных платформах с целью устранения его внезапного изменения при смене аудиофайлов, что обеспечивает комфортное восприятие звука [10].

Обзор моделей средства распознавания речи Whisper

Whisper – модель преобразования речи в текст, являющаяся открытой и основанной на применении нейронных сетей – трансформеров. В ней реализованы две ключевые составляющие: кодер и декодер. Кодер обрабатывает акустический сигнал, разбивая его на фрагменты и преобразуя в спектрограммы. Декодер на основе этих преобразований генерирует текстовую расшифровку (наиболее вероятную последовательность символов или слов, соответствующих речи) в зависимости от языка исходного речевого сообщения [11].

Наиболее распространенные варианты моделей Whisper, отличающиеся друг от друга числом параметров и структурой, влияю-

щих на их производительность, точность и требования к вычислительным ресурсам, представлены в табл. 2.

Модели Tiny, Base, Small, Medium обучены на 680 тыс. часах аудиозаписей, представленных сообщениями на 97 языках с посторонними шумами, смехом и другими звуками. Модель Large-v3 была обучена на смеси примерно 1 миллиона часов размеченного и 4 миллионов часов псевдоразмеченных аудиоданных. Входные данные включали аудиозаписи на множестве языков и в различных условиях записи. Модель Large-v3-turbo также обучалась на 5 миллионах часов размеченного аудио, что обеспечивает её способность точно распознавать речь на более чем 99 языках, включая шумные и сложные аудиозаписи. За счёт архитектурных изменений (сокращение числа слоев декодера с 32 до 4) модель стала легче и быстрее, сохранив при этом качество.

Выбор конкретной реализации модели Whisper определяется преимущественно типом решаемой задачи и доступными вычислительными ресурсами. Следовательно, при выборе модели важно учитывать баланс требований по точности распознавания и вычислительных ресурсов, а также условия формирования речевых сообщений (наличие посторонних шумов).

Экспериментальная оценка точности распознавания речи моделью искусственного интеллекта Whisper при проведении социологических опросов (результаты исследования)

В исследовании использовалась социологическая анкета из 30 вопросов, в т.ч. один вопрос предусматривал множественный выбор ответов. Анализируемый аудиофайл представлял собой запись ответов респондентов на вопросы анкеты, выполненную микрофоном устройства Samsung S20. Запись производилась с расстояния 3 – 4 см от микрофона обычной разговорной громкостью голоса. Среднее значение громкости речи при измерении составило –18 LUFS (для измерения используется программа Adobe Audition 2023 с установленным стандартом EBU R 128 и глубиной кодирования файла 16 бит), что соответствует громкости речи 50 дБ.

Для оценки качества распознавания речи в начале распознавание происходит на аудиоданных, характеризующихся отсутствием фонового шума или его уровнем, существенно ниже уровня громкости речи, а в конце – на аудиоданных, в которых искусственно внесены акустические шумы различной громкости.

Данные средства распознавания речи будем сравнивать с помощью показателя флективной частоты ошибки слов – IWER (Inflectional Word Error Rate).

$$IWER = \frac{S_{hard}*C_{hard}+S_{soft}*C_{soft}+D+I}{N},$$

где C_{hard} – вес «грубой» ошибки, приведшей к замене изначального слова другим отличающимся по смыслу словом (например, диаграмма – дейтаграмма); S_{hard} – количество «грубых» ошибок; C_{soft} – вес «негрубой» ошибки, не повлиявшей на основу слова (например, сделано – сделана); S_{soft} – количество «негрубых» ошибок; D – количество удаленных слов; I – количество излишне добавленных слов; N – общее количество слов в эталоне (речи) [12].

По определению, $C_{soft}<C_{hard}$, $C_{soft}<1$, $C_{hard}>=1$, при этом в данном исследовании принято $C_{soft} = 0,5$, $C_{hard} = 1$.

Исследование предполагает распознавание идентичного аудиофайла с использованием каждой модели, после чего вычисляется флективная частота ошибок слов. На основании полученных показателей выполняется выбор модели с наибольшей точностью распознавания.

Количество слов в эталонной фразе ответов респондентов N = 257 слов. Например, расчет IWER для Large-v3-turbo согласно (1):

IWER = (6·0,5 + 3)/257 = 0,0233 (2,33%).

Результаты эксперимента представлены в табл. 3.

После произведенной оценки распознавания аудиофайла без фонового шума каждой моделью Whisper, можно прийти к выводу, что large-v3-turbo лучше всех распознала данный аудиофайл.

В дальнейшем необходимо проверить каждую модель на основе зашумленной речи, т.е. в изначальной речи содержится шум улицы.

Громкость речи: –18 LUFS

1) У дороги: –27,9 LUFS
2) 10 метров от дороги: –29,6 LUFS
3) 15 метров от дороги: –30,7 LUFS
4) 20 метров от дороги: –36 LUFS
5) Во дворе жилого сектора: –40 LUFS

Шум добавлялся методом линейного микширования с коэффициентом β, рассчитанным по формуле $\beta = 10^{(L\_шум - L\_речь)/20}$, где

Таблица 3

Точность распознавания речи при отсутствии шумов открытого пространства

	Удалено слов	Грубые ошибки	Негрубые ошибки	Излишне добавлено слов	IWER, %
Large-v3-turbo	0	0	6	3	2,33
Large-v3	1	4	11	2	4,86
Medium	0	8	9	1	5,25
Small	1	21	16	3	12,84
Base	19	37	26	5	28,79
Tiny	1	41	25	14	26,65

L_речь = –18 LUFS, L_шум варьировалось от –27,9 до –40 LUFS в зависимости от сценария. После микширования сигнал нормировался к –18 LUFS по стандарту EBU R 128 [10]. Все операции выполнены в Adobe Audition 2023.

Значения параметра IWER в зависимости от условий для каждого средства распознавания представлены в табл. 4 – 8.

Значения параметра IWER в зависимости от условий для различных моделей представлены в табл. 9.

Для удобства визуального анализа таблицы 9 представим полученные данные в виде графиков (рис. 1), из которых следует, что наилучшие результаты продемонстрировала модель Large-v3-turbo.

Таблица 4

Точность распознавания речи при нахождении непосредственно около источника посторонних шумов (дороги)

	Разность громкостей речи и шума, LUFS	Удалено слов	Грубые ошибки	Негрубые ошибки	Излишне добавлено слов	IWER, %
Large-v3-turbo	9,9	2	7	7	0	4,86
Large-v3	9,9	2	6	16	5	8,17
Medium	9,9	12	8	25	2	13,42
Small	9,9	3	29	18	3	17,12
Base	9,9	9	60	38	8	37,35
Tiny	9,9	15	71	25	11	42,61

Таблица 5

Точность распознавания речи при нахождении на расстоянии 10 м от источника посторонних шумов (дороги)

	Разность громкостей речи и шума, LUFS	Удалено слов	Грубые ошибки	Негрубые ошибки	Излишне добавлено слов	IWER, %
Large-v3-turbo	11,6	1	2	5	1	2,53
Large-v3	11,6	2	4	16	1	5,84
Medium	11,6	3	1	15	0	4,47
Small	11,6	3	14	13	1	9,53
Base	11,6	1	48	22	2	24,12
Tiny	11,6	5	56	21	4	29,38

Таблица 6

Точность распознавания речи при нахождении на расстоянии 15 м от источника посторонних шумов (дороги)

	Разность громкостей речи и шума, LUFS	Удалено слов	Грубые ошибки	Негрубые ошибка	Излишне добавлено слов	IWER, %
Large-v3-turbo	12,7	1	2	4	0	1,95
Large-v3	12,7	1	4	10	1	4,28
Medium	12,7	0	13	5	0	6,03
Small	12,7	2	16	17	2	11,09
Base	12,7	7	43	15	6	24,71
Tiny	12,7	10	57	23	3	31,71

Таблица 7

Точность распознавания речи при нахождении на расстоянии 20 м от источника посторонних шумов (дороги)

	Разность громкостей речи и шума, LUFS	Удалено слов	Грубые ошибки	Негрубые ошибка	Излишне добавлено слов	IWER, %
Large-v3-turbo	18	1	4	6	0	2,14
Large-v3	18	1	2	11	1	3,7
Medium	18	1	11	3	0	5,25
Small	18	4	11	20	0	9,73
Base	18	2	30	22	2	17,51
Tiny	18	4	50	27	3	27,43

Таблица 8

Точность распознавания речи при нахождении в жилом секторе при отсутствии движения транспортных средств

	Разность громкостей речи и шума, LUFS	Удалено слов	Грубые ошибки	Негрубые ошибка	Излишне добавлено слов	IWER, %
Large-v3-turbo	22	1	0	2	0	0,79
Large-v3	22	4	0	1	0	1,75
Medium	22	1	2	9	1	3,31
Small	22	3	16	11	0	9,53
Base	22	0	23	19	1	13,06
Tiny	22	7	40	19	5	23,93

Таблица 9

Сводная таблица значений параметра IWER для моделей, %

	У дороги	10 метров от дороги	15 метров от дороги	20 метров от дороги	Во дворе
Large-v3-turbo	4,86	2,53	1,95	2,14	0,79
Large-v3	8,17	5,84	4,28	3,7	1,75
Medium	13,42	4,47	6,03	5,25	3,31
Small	17,12	9,53	11,09	9,73	9,53
Base	37,35	24,12	24,71	17,51	13,06
Tiny	42,61	29,38	31,71	27,43	23,93

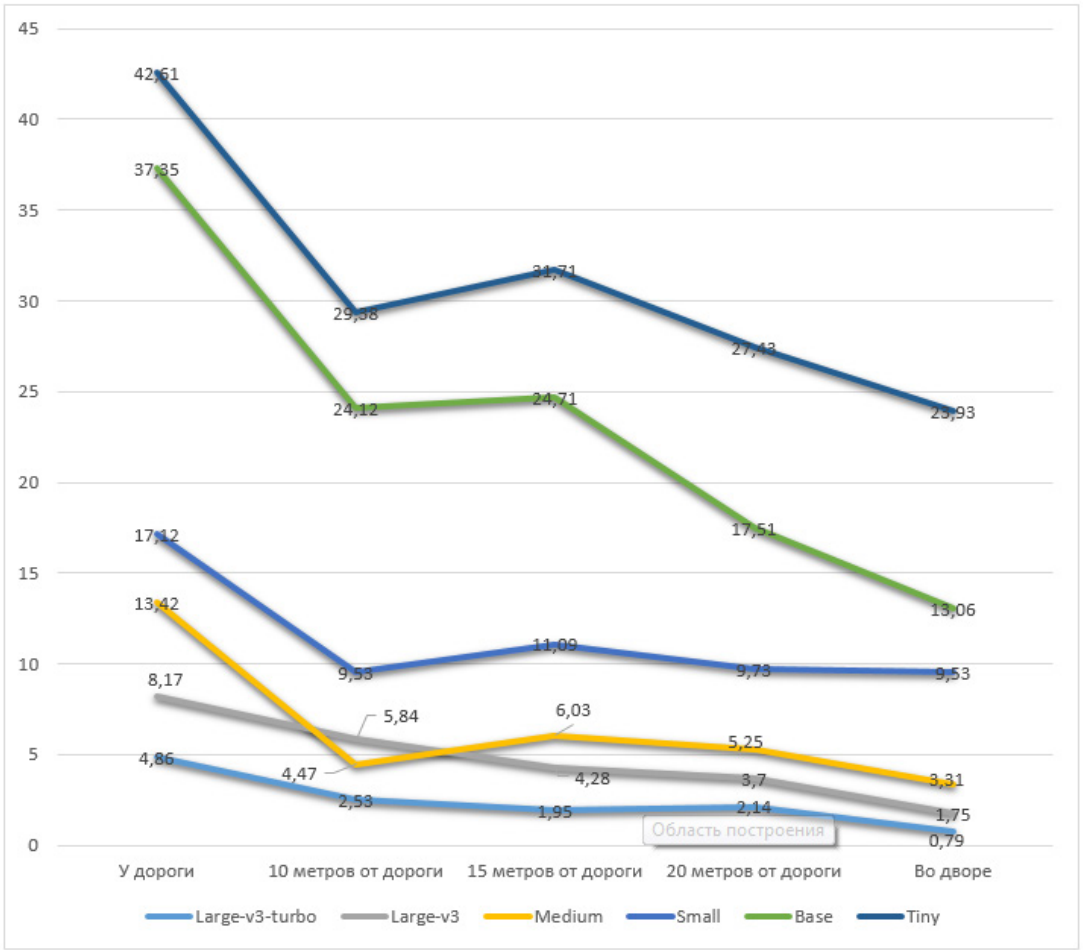

Рис. 1. Зависимости обобщенного показателя качества распознавания IWER от места проведения социологического опроса

Особенность распознавания речи у моделей Medium, Small, Base и Tiny проявляется в следующем: количество флективных ошибок при расстоянии 10 метров от дороги меньше, чем при расстоянии 15 метров. Предположительно, данное явление обусловлено особенностями обучения этих моделей.

Кроме того, установлено, что показатели качества распознавания в условиях жилого сектора (при отсутствии интенсивных акустических помех) превышают значения, полученные при тестировании на аудиофайлах без фонового шума. Данное явление может свидетельствовать о том, что обучающая вы-

борка моделей включала аудиоданные с низким уровнем фоновых шумов, что способствовало формированию устойчивости алгоритмов к подобным акустическим условиям.

Заключение

Экспериментальная оценка подтвердила, что для автоматизации транскрибации интервью в условиях нестационарного шума оптимальным выбором является модель Large-v3-turbo. Она обеспечивает снижение показателя IWER до 0,79% в благоприятных условиях и сохраняет работоспособность при соотношении сигнал/шум около 10 дБ. В отличие от облегченных версий, данная модель реже допускает грубые смысловые ошибки, что критично для последующего анализа социологических данных. Таким образом, внедрение указанной модели в специализированные информационные системы позволит сократить трудозатраты на обработку речевых ответов без существенной потери качества.

Перспективы исследований

Развитие этого направления исследований находится в адаптации алгоритмов под предметную область. Стандартные обучающие выборки Whisper не учитывают специфику социологической лексики, поэтому логичным шагом станет дообучение (fine-tuning) на профильных данных. Также планируется оценить эффективность гибридных конвейеров, где распознавание дополняется семантической коррекцией через языковые модели типа ruGPT-3. Нельзя исключать и необходимость тестирования на языках народов РФ для расширения географии исследований. Наконец, для полевых работ актуальна задача оптимизации моделей под мобильные платформы – сравнение квантованных версий поможет найти баланс между скоростью работы и точностью распознавания.

Литература

1. Садыкова А. А., Амиргалиев Е. Н. Изучение применения автоматического распознавания речи // Colloquium-journal. 2020. № 11 (63). С. 44–47. – DOI: 10.24411/2520-6990-2020-11728.

2. Ворошилова А. И. Искусственный интеллект в социологическом исследовании: опыт использования для обработки и анализа интервью // Социодинамика. 2025. № 2. – Электрон. ресурс. – URL: https://nbpublish.com/library_read_article.php?id=73330 (дата обращения: 26.04.2026).

3. Гапочкин А. В. Нейронные сети в системах распознавания речи // Science Time. 2014. № 11. С. 483–492.

4. Мырадов М.Т., Назарова С.Н. Современные технологии распознавания речи // Наука и мировоззрение. 2024. № 3. С. 45–52.

5. Рахманов М. А. Сравнительный обзор современных моделей распознавания речи // Международная научно-практическая конференция «Smart Technologies and Innovative Approaches in Research: Building a Digital Future»: сборник материалов. 2025. 22 февр. С. 572–577. – DOI: 10.5281/zenodo.14918743.

6. Graham C., Roll N. Evaluating OpenAI's Whisper ASR: Performance analysis across diverse accents and speaker traits // Journal of the Acoustical Society of America. 2024. Vol. 155, Iss. 2. P. 1123–1135. PMID: 38391582. – DOI: 10.1121/10.0024876.

7. Andreyev A. Quantization for OpenAI's Whisper Models: A Comparative Analysis // arXiv preprint arXiv:2503.09905 [cs.SD]. 2025. – DOI: 10.48550/arXiv.2503.09905.

8. Ковригина Е. А. Компьютер и распознавание речи // Вестник компьютерных и информационных технологий. 2021. № 4. С. 22–29.

9. Физика: акустика. Громкость и интенсивность звука: учеб. пособие / под ред. В.Н. Петрова. М.: Физматлит, 2019. 184 с.

10. EBU Tech Doc 3343-2011v2. Practical Guidelines for Production & Implementation of EBU R 128. Geneva: European Broadcasting Union, 2011. 48 p. – URL: https://tech.ebu.ch/docs/tech/tech3343.pdf (дата обращения: 26.04.2026).

11. OpenAI. Whisper: Robust Speech Recognition via Large-Scale Weak Supervision // arXiv preprint arXiv:2212.04356 [cs.LG]. 2022. – DOI: 10.48550/arXiv.2212.04356.

12. Бовбель Е. И., Паршин В. В. Нейронные сети в системах автоматического распознавания речи // Зарубежная радиоэлектроника. 1998. № 4. С. 49–65.

References

1. Sadykova A. A., Amirgaliev E. N. Izuchenie primeneniya avtomaticheskogo raspoznavaniya rechi // Colloquium-journal. 2020. № 11 (63). P. 44–47. – DOI: 10.24411/2520-6990-2020-11728.

2. Voroshilova A. I. Iskusstvennyy intellekt v sotsiologicheskom issledovanii: opyt ispol'zovaniya dlya

obrabotki i analiza interv'yu // Sotsiodinamika. 2025. № 2. – Electronic resource. – URL: https://nbpublish.com/library_read_article.php?id=73330 (date of access: 26.04.2026).

3. Gapochkin A. V. Neyronnye seti v sistemakh raspoznavaniya rechi // Science Time. 2014. № 11. P. 483–492.

4. Myradov M. T., Nazarova S. N. Sovremennye tekhnologii raspoznavaniya rechi // Nauka i mirovozzrenie. 2024. № 3. P. 45–52.

5. Rakhmanov M. A. Sravnitel'nyy obzor sovremennykh modeley raspoznavaniya rechi // International Scientific and Practical Conference "Smart Technologies and Innovative Approaches in Research: Building a Digital Future": proceedings. 2025. 22 Feb. P. 572–577. – DOI: 10.5281/zenodo.14918743.

6. Graham C., Roll N. Evaluating OpenAI's Whisper ASR: Performance analysis across diverse accents and speaker traits // Journal of the Acoustical Society of America. 2024. Vol. 155, Iss. 2. P. 1123–1135. PMID: 38391582. – DOI: 10.1121/10.0024876.

7. Andreyev A. Quantization for OpenAI's Whisper Models: A Comparative Analysis // arXiv preprint arXiv:2503.09905 [cs.SD]. 2025. – DOI: 10.48550/arXiv.2503.09905.

8. Kovrigina E. A. Komp'yuter i raspoznavanie rechi // Vestnik komp'yuternykh i informatsionnykh tekhnologiy. 2021. № 4. P. 22–29.

9. Fizika: akustika. Gromkost' i intensivnost' zvuka: ucheb. posobie / pod red. V. N. Petrova. M.: Fizmatlit, 2019. 184 p.

10. EBU Tech Doc 3343-2011v2. Practical Guidelines for Production & Implementation of EBU R 128. Geneva: European Broadcasting Union, 2011. 48 p. – URL: https://tech.ebu.ch/docs/tech/tech3343.pdf (date of access: 26.04.2026).

11. OpenAI. Whisper: Robust Speech Recognition via Large-Scale Weak Supervision // arXiv preprint arXiv:2212.04356 [cs.LG]. 2022. – DOI: 10.48550/arXiv.2212.04356.

12. Bovbel' E. I., Parshin V. V. Neyronnye seti v sistemakh avtomaticheskogo raspoznavaniya rechi // Zarubezhnaya radioelektronika. 1998. № 4. P. 49–65.

АРАКЧАА Айдаш Мергенович, сотрудник, Академия ФСО России, г. Орёл. E-mail: arakchaaaydash003@gmail.com

ДЕМИН Дмитрий Михайлович, сотрудник, Академия ФСО России, г. Орёл. E-mail: d1meplg@yandex.ru

ЖУСОВ Дмитрий Леонидович, кандидат технических наук, доцент, сотрудник, Академия ФСО России, г. Орёл. E-mail: d.zhusov@mail.ru

МАКЕЕВ Сергей Михайлович, кандидат технических наук, доцент кафедры компьютерной безопасности и прикладной алгебры, Челябинский государственный университет, 454001, г. Челябинск, ул. Братьев Кашириных, 129.; доцент кафедры защиты информации, Южно-Уральский государственный университет (национальный исследовательский университет), 454080, г. Челябинск, пр. Ленина, 76. E-mail: maksm57@yandex.ru

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

ARAKCHAA Aidash Mergenovich, Staff Member, Academy of the Federal Guard Service of Russia, Orel. E-mail: arakchaaaydash003@gmail.com

DEMIN Dmitry Mikhailovich, Staff Member, Academy of the Federal Guard Service of Russia, Orel. E-mail: d1meplg@yandex.ru

ZHUSOV Dmitry Leonidovich, PhD in Technical Sciences, Associate Professor, Staff Member, Academy of the Federal Guard Service of Russia, Orel. E-mail: d.zhusov@mail.ru

MAKEEV Sergey Mikhailovich, PhD in Technical Sciences, Associate Professor at the Department of Computer Security and Applied Algebra, Chelyabinsk State University, 129 Bratiev Kashirinykh St., Chelyabinsk 454001, Russia; Associate Professor at the Department of Information Security, South Ural State University (National Research University), 76 Lenin Ave., Chelyabinsk 454080, Russia. E-mail: maksm57@yandex.ru

SOKOLOV Alexander Nikolaevich, PhD in Technical Sciences, Associate Professor, Head of the Department of Information Security, South Ural State University (National Research University), 76 Lenin Ave., Chelyabinsk 454080, Russia. E-mail: sokolovan@susu.ru

УДК 004.72:65.011.56 Вестник УрФО № 2(60) / 2026, с. 24–31

Быкасов А. В., Соколов А. Н.

DOI: 10.14529/secur260203

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ВЛИЯНИЯ АКТИВНОГО СКАНИРОВАНИЯ НА СТАБИЛЬНОСТЬ ФУНКЦИОНИРОВАНИЯ СЕТИ АСУ ТП

В статье рассматривается задача обеспечения кибербезопасности автоматизированных систем управления технологическими процессами (АСУ ТП) при применении методов активного сетевого сканирования. Предложена аналитическая математическая модель трафика промышленной сети, построенная на основе теории массового обслуживания (ТМО). Модель позволяет прогнозировать задержки при передаче данных, длину очередей сообщений и коэффициент загрузки сетевых узлов до начала проведения сканирования. Показана взаимодополняемость аналитического подхода и современных моделей доверия к методам активного сканирования. Сформулированы критерии безопасного внедрения сканирования, обеспечивающие непрерывность технологического процесса и ограничение суммарной задержки $D_{total} < 10$ мс. Установлено, что поддержание коэффициента загрузки $\rho \leq 0{,}7$ предотвращает экспоненциальный рост времени ожидания в очередях. Результаты исследования могут быть применены при проектировании регламентов кибербезопасности, планировании окон обслуживания и разработке адаптивных алгоритмов сканирования.

Ключевые слова: *АСУ ТП, теория массового обслуживания, активное сканирование, математическое моделирование, сетевой трафик, кибербезопасность, задержки передачи, коэффициент загрузки, модель доверия.*

MATHEMATICAL MODELING OF THE IMPACT OF ACTIVE SCANNING ON THE STABILITY OF THE INDUSTRIAL CONTROL NETWORK

This paper addresses the challenge of ensuring the cybersecurity of industrial control systems (ICS) when employing active scanning methods. An analytical mathematical model of industrial network traffic based on queuing theory is proposed. This model allows for prediction of data transmission delays, message queue lengths, and network node utilization rates prior to the commencement of scanning operations. The complementarity between this analytical approach and modern trust models regarding active scanning methods is demonstrated. Criteria for the safe implementation of network scanning are formulated, ensuring both the continuity of technological processes and a total delay limit of $D_{total} < 10$ ms. It is established that maintaining a utilization rate of $\rho \leq 0.7$ prevents the exponential growth of queuing wait times. The findings of this study can be applied in the design of cybersecurity protocols, the scheduling of maintenance windows, and the development of adaptive scanning algorithms.

Keywords: *ICS, queuing theory, active scanning, mathematical modeling, network traffic, cybersecurity, transmission delays, load factor, trust model.*

Обеспечение информационной безопасности автоматизированных систем управления технологическими процессами (АСУ ТП) является важной задачей для объектов критической информационной инфраструктуры (КИИ), особенно в условиях нарастающей волны целевых атак на промышленные системы. Как следствие наличия модернизируемой базы кибератак, постоянно возрастает потребность в проведении регулярной оценки уязвимостей сетевой инфраструктуры. Активное сканирование портов и служб попрежнему остается одним из самых эффективных способов выявления слабых мест, но в АСУ ТП его целесообразность подвергается сомнению по причине наличия огромного количества серьезных рисков. Избыточные запросы могут привести к перегрузке каналов связи и увеличению задержек передачи данных, а также вызвать сбои в работе программируемых логических контроллеров (ПЛК) и человеко-машинных интерфейсов (HMI).

Существующие подходы к оценке воздействия сканирования носили эмпирический характер и не учитывали возможности обеспечения необходимой надежности и безопасности результатов. В работе [1] рассматриваются основные методы активного сканирования сетей АСУ ТП и описываются алгоритмы их реализации. В результате применения методов активного сканирования была построена UML-диаграмма активов сети АСУ ТП. Сформулированы условия доверия, применимые к методам активного сканирования, которые позволят разрабатывать безопасные методы сбора и анализа сетевой информации с целью создания верифицированной системы активного сканирования для поиска уязвимостей в сетях АСУ ТП.

В работе [2] представлена статистическая модель оценки влияния комбинированного метода активного сканирования на стабильность функционирования сети АСУ ТП. Проведены эксперименты для оценки влияния активного сканирования на стабильность функционирования промышленной сети. Полученные результаты, представленные в

сравнении с результатами для распространенного инструмента сетевого сканирования Nmap, дают сделать вывод о том, что комбинированный метод активного сканирования оказывает меньшее влияние на сеть, а также обеспечивает стабильность при сканировании, тем самым не нарушая работоспособность сети АСУ ТП.

Несмотря на высокую ценность моделей доверия и статистической валидации, для продуктивного планирования окон обслуживания и нормативного регулирования интенсивности сканирования требуется аналитический аппарат. В данной работе предлагается применение теории массового обслуживания (ТМО), фундаментальные положения которой изложены в классических работах [3, 4], для построения аналитической модели трафика АСУ ТП. При выборе параметров модели учтены рекомендации по архитектуре защищенных сетей, приведенные в [7, 8], а также требования к защите информации в промышленных сетях, регламентированные в [9] и нормативных документах ФСТЭК России [10 - 12]. Проблемы комплексного обеспечения безопасности технологических сетей, отмеченные в [13], также учитываются при формировании критериев устойчивости.

Для построения математической модели рассмотрим типовую архитектуру сети промышленной системы управления со звездообразной топологией (рис. 1).

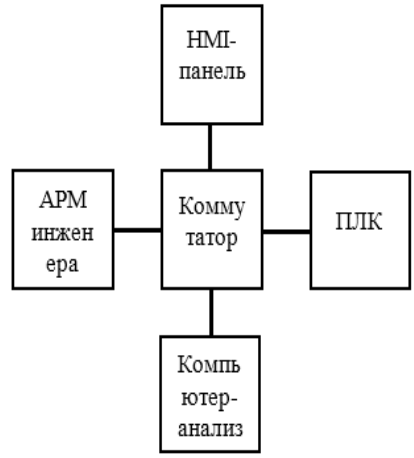

Рис. 1. Архитектура сети АСУ ТП со звездообразной топологией

Схема (рис. 1) включает следующие компоненты: HMI-панель (генерация трафика визуализации ~100 пак./с), ПЛК (обработка сенсорных данных ~50 пак./с), центральный промышленный коммутатор (пропускная способность портов 1 Гбит/с), АРМ инженера по безопасности (узел диагностики и сканирования) и сервер истории данных. Соединение осуществляется по технологии Ethernet с использованием протоколов Modbus TCP, PROFINET, Ethernet/IP, требования к безопасности которых регламентированы в [8, 9].

Трафик в сетях АСУ ТП имеет смешанный характер и делится на три основных типа, характеристики которых представлены в табл. 1.

Таблица 1

Характеристики типов трафика в сети АСУ ТП

Тип трафика	Интенсивность, пак./с	Приоритет	Распределение интервалов
Периодический (ПЛК → HMI)	50–100	Высокий	Равномерное
Событийный (аварии, алерты)	1–10 (спорадически)	Критический	Экспоненциальное
Диагностический и сканирующий	20–500 (управляемый)	Низкий	Пуассоновское

Периодический трафик характеризуется регулярной передачей данных с фиксированным интервалом $T = 1$с. Для его описания используется равномерное распределение:

$$f(t) = \begin{cases} \frac{1}{b-a}, & a \le t \le b \\ 0, & \text{иначе} \end{cases}, \qquad (1)$$

где – время ожидания между событиями,

[a, b] – интервал времени, в течение которого происходит передача пакета.

Событийный трафик возникает при аварийных ситуациях. Моделируется экспоненциальным распределением:

$$f(t) = \lambda e^{-\lambda t},\ t \ge 0, \qquad (2)$$

где t – время ожидания между аварийными событиями,

λ – интенсивность аварийных событий (среднее количество событий в единицу времени).

Диагностический и сканирующий трафик генерируется АРМ инженера и системами мониторинга сети. Характеризуется пуассонов-

ским потоком с интенсивностью , которая является управляемым параметром.

Математическая модель сетевого трафика на основе ТМО

Центральным элементом сети является коммутатор, который моделируется как система с очередями (СМО) (рис. 2). Для описания процессов поступления и обработки пакетов применяется нотация Кендалла A/B/c/K/N/D [5], где A – характер входящего потока (например, M – пуассоновский поток), B – распределение времени обслуживания одного пакета (например, M – экспоненциальное), c – количество каналов обслуживания, K – буфер очереди пакетов, N – общее количество пакетов (если очередь пакетов бесконечна, параметр опускается), D – тип обслуживания очереди пакетов (например, FIFO – первый пришел, первый обслужен).

Для базовой модели принята M/M/1,

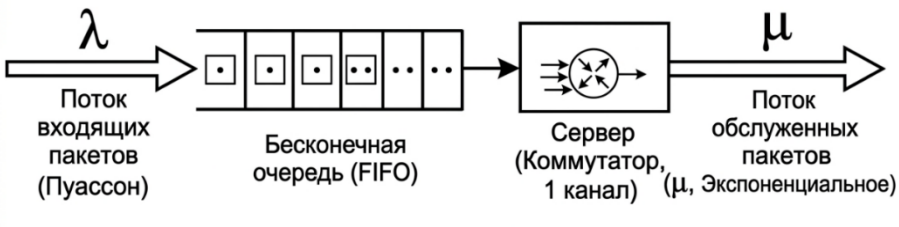

Рис. 2. Модель СМО M/M/1 для коммутатора сети АСУ ТП

предполагающая пуассоновский вход, экспоненциальное обслуживание и один канал.

Коэффициент загрузки системы:

$$\rho = \frac{\lambda}{\mu}, \quad (3)$$

где λ – интенсивность входящего потока пакетов (среднее количество пакетов в единицу времени),

μ – интенсивность обслуживания пакетов (среднее количество пакетов, которое канал может обслужить в единицу времени).

Условие стабильности: ρ < 1. Для сетей реального времени рекомендуется ρ ≤ 0,7 [6, 7].

Среднее время ожидания в очереди:

$$W_q = \frac{\rho}{\mu(1-\rho)} = \frac{\lambda}{\mu(\mu-\lambda)}. \quad (4)$$

Среднее число заявок в системе:

$$L_s = \frac{\rho}{1-\rho}. \quad (5)$$

Для АСУ ТП критически важно обеспечение приоритетной обработки технологического трафика. В модели M/M/1 с двумя классами приоритетов среднее время ожидания для технологического (высший приоритет) и сканирующего (низший приоритет) трафика определяется соответственно формулами:

$$W_q^{(1)} = \frac{1}{\mu(1-\rho_1)}, \quad (6)$$

$$W_q^{(2)} = \frac{1}{\mu(1-\rho_1)(1-\rho_1-\rho_2)}, \quad (7)$$

где $\rho_1 = \frac{\lambda_{tech}}{\mu}, \rho_2 = \frac{\lambda_{scan}}{\mu}$.

Формула (7) демонстрирует, что при приближении $\rho_1 + \rho_2$ к 1, задержки сканирующего трафика возрастают экспоненциально, не влияя на критический технологический обмен.

Расчетная модель задержек и критерии безопасности

Общая задержка передачи пакета складывается из нескольких компонент [4, 5]:

$$D_{total} = D_{trans} + D_{prop} + D_{proc} + W_q, \quad (8)$$

где D_{trans} = L/R – задержка передачи, D_{prop} = = d/v – задержка распространения, D_{proc} – задержка обработки коммутатором (10–50 мкс), W_q – очередная задержка. Требования к максимальным задержкам в промышленных сетях (D_{total} < 10 мс) соответствуют рекомендациям [8, 9].

Результаты расчета параметров сети при различной интенсивности сканирования представлены в табл. 2.

По графику можно заметить, что зависимость W_q от ρ при ρ > 0,7 имеет нелинейный характер (рис. 3), причем найденные пороговые значения кореллируют с результатами статистического моделирования [2], где экспериментально показано, что комбинированный метод активного сканирования выдает выборочное математическое ожидание для среднего показателя всплесков трафика на порядок ниже, чем метод Nmap. Это доказывает, что при условии ρ ≤ 0,7 и использовании комбинированных методов активного сканирования гарантируется предсказуемость поведения сети и отсутствие критических перегрузок.

Разработаны три критерия на основании условий доверия из [1] и нормативной базы [10 - 12], с помощью которых можно сформулировать требований безопасности для методов активного сканирования:

Таблица 2

Параметры сети при различной интенсивности сканирования

λ_{scan}, пак./с	ρ	W_q, мкс	D_{total}, мс	Статус
0	0,006	0,024	0,151	Безопасно
5000	0,206	1,03	0,116	Безопасно
10000	0,406	2,71	0,128	Безопасно
17350	0,700	9,33	0,244	Граница (ρ = 0,7)
22000	0,886	20,75	0,147	Опасно

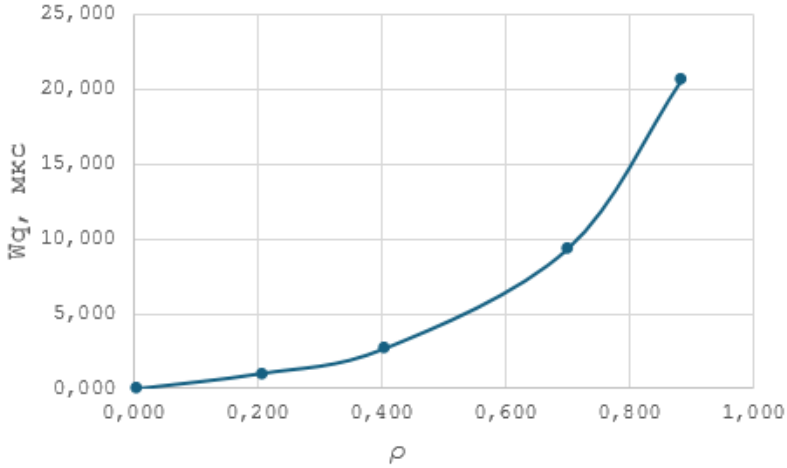

Рис. 3. Зависимость среднего времени ожидания от коэффициента загрузки

1. Непрерывность работы: вероятность отказа системы с ограниченным размером буфера очереди пакетов K, согласно (11):

$$P_{\text{отказа}} = \frac{(1-\rho)\rho^K}{1-\rho^{K+1}} < 10^{-6}, \quad (9)$$

при K = 100, что соответствует требованиям надежности промышленных сетей.

2. Минимизация нагрузки: в соответствии с первым условием доверия [1], метод активного сканирования должен исключить передачу избыточных данных:

$$\lambda_{\text{tech}} + 0_{\text{scan}} \leq 0{,}7 \cdot \mu. \quad (10)$$

3. Контроль нагрузки на устройства:

$$U_{PLC} = \frac{\lambda_{in}+\lambda_{out}}{f_{CPU}} \times 100\% < 70\%. \quad (11)$$

4. Интегральный показатель безопасности сети формируется по формуле:

$$S = w_1 \cdot \frac{D_{max} - D_{total}}{D_{max}} + w_2 \cdot \frac{0{,}7-\rho}{0{,}7} + w_3 \cdot (1 - P_{\text{отказа}}), \; S \geq 0{,}7. \quad (12)$$

В соответствии с моделью доверия, представленной в [1], для получения безопасного взаимодействия методов активного сканирования с сетями АСУ ТП необходимо выполнение трех условий доверия:

1. Исключение избыточного трафика: метод не должен создавать стабильный избыточный трафик объемом более половины минимальной полосы пропускания устройства. В терминах ТМО это условие формулируется как ограничение на интенсивность сканирующего потока:

$$\lambda_{scan} \leq \frac{Cmin}{2L_{avg}}, \quad (13)$$

где C_{min} – минимальная пропускная способность канала, L_{avg} – средний размер пакета.

2. Эффективное построение топологии сети: методология должна иметь алгоритм, позволяющий идентифицировать все включенные в сеть узлы. В соответствии с моделью ТМО условием достижения полноты построения топологической карты сети при минимальной интенсивности сканирования будет требование формулы определения вероятности

$$P_{\text{detect}} = 1 - e^{-\lambda\text{scan} \cdot \text{Tscan}} \geq 0{,}99, \quad (14)$$

которой удовлетворяет время T_{scan}, отводимое на сканирование сегмента сети.

3. Совместимость с устройствами АСУ ТП: методология должна подбирать запрос, подходящий для конкретного узла и не нарушающий его работоспособность. Реализация условия обеспечивается приоритизацией тра-

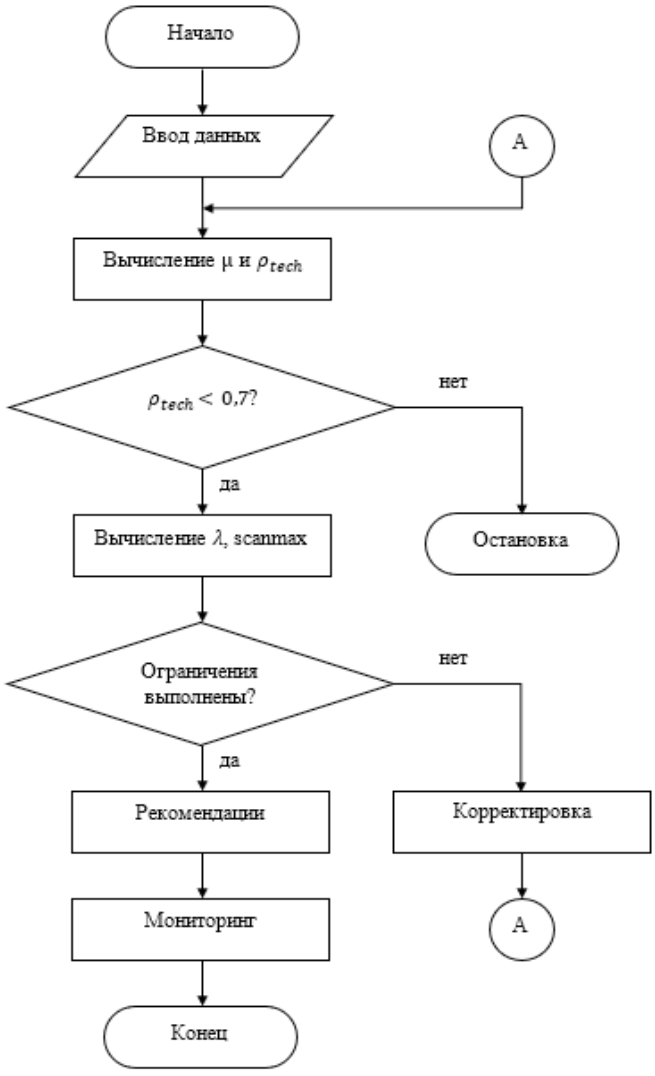

Рис. 4. Алгоритм расчета безопасных параметров сканирования

фика и адаптивным контролем интенсивности сканирования:

$$\lambda_{scan}(t) = \begin{cases} \lambda_{min}, & \rho < 0{,}5 \\ \lambda_{min} + k \cdot (0{,}7 - \rho), & 0{,}5 \le \rho \le 0{,}7 \\ 0, & \rho > 0{,}7 \end{cases} \quad (15)$$

Алгоритм расчета безопасных параметров сканирования (рис. 4) использует следующие шаги:

1) Сбор входных параметров (λ_{tech}, C, D_{max}).

2) Расчет пропускной способности $\mu = \frac{C}{L_{avg}}$ и коэффициента загруженности ρ_{tech}, проверка условия $\rho_{tech} < 0{,}7$.

3) Определение резерва по интенсивности сканирования $\lambda_{scan}^{max} = (0{,}7 - \rho_{tech})\mu$.

4) Проверка ограничений по среднему времени ожидания W_q и общей задержки передачи пакета D_{total} с использованием формул (4)–(8).

5) Формирование рекомендаций и временных окон для использования активного сканирования.

6) Мониторинг промышленной сети в реальном времени с адаптивным корректированием интенсивности трафика сканирования λ_{scan}.

Использование теории массового обслуживания для моделирования трафика АСУ ТП предоставляет строгую математическую базу для оценивания безопасности методов активного сетевого сканирования. Выявлено, что выполнение $\rho \le 0{,}7$ и $D_{total} < 10$ мс гарантирует сохранение непрерывности сетевого взаимодействия устройств промышленной сети. Аналитические результаты согласуются с современными моделями доверия [1] и статистическими моделями оценки влияния активного сканирования [2], которые подтверждают, что комбинированные методы

Таблица 3

Уровни интенсивности мониторинга промышленной сети и действия при превышении порогов

Уровень	p	D_{total}	Действие
Нормальный	<0,5	<1 мс	Продолжение сканирования
Предупреждение	0,5–0,7	1–5 мс	Снижение λ_{scan} на 25%
Тревога	0,7–0,85	5–10 мс	Снижение λ_{scan} на 50%
Критический	> 0,85	> 10 мс	Аварийная остановка сканирования

активного сканирования с управляемой интенсивностью гарантируют стабильность сети АСУ ТП при минимальных задержках внутри сети.

Разработанный алгоритм (рис. 4) и интегральный критерий безопасности (формула 12) могут быть внедрены в системы управления инцидентами (SIEM) и в оркестраторы безопасности (SOAR) для автоматизированного планирования окон анализа уязвимостей. Имеющимися перспективами исследований являются экспериментальная верификация модели на тестовых стендах с реальным промышленным оборудованием, учет приоритетной дисциплины обслуживания (Priority Queuing) и интеграция с имитационным моделированием (OMNeT++, NS-3) для верификации применяемых методов активного сканирования в условиях гетерогенных промышленных протоколов.

Литература

1. Bykasov A. V., Sokolov A. N., Barinov A. E. Trust Conditions for Active Scanning Methods for Finding Vulnerabilities in ICS Networks // 2023 International Russian Automation Conference (RusAutoCon). IEEE, 2023. DOI: 10.1109/RusAutoCon58002.2023.10272827.

2. Быкасов А. В., Соколов А. Н. Статистическая модель оценки влияния комбинированного метода активного сканирования на стабильность функционирования сети АСУ ТП // Вестник УрФО. Безопасность в информационной сфере. 2024. № 2(52). С. 68–76. DOI: 10.14529/secur240207.

3. Клейнрок Л. Теория массового обслуживания. – М.: Машиностроение, 1979. – 432 с.

4. Бронштейн О.И., Духовный И.М. Модели приоритетного обслуживания в информационно-вычислительных системах. – М.: Наука, 1976. – 220 с.

5. Гнеденко Б.В., Коваленко И.Н. Введение в теорию массового обслуживания. – 2-е изд. – М.: Наука, 1987. – 336 с.

6. Bertsekas D., Gallager R. Data Networks. – 2nd ed. – Prentice Hall, 1992. – 640 p.

7. Столлингс В. Современные компьютерные сети. – 4-е изд. – СПб.: Питер, 2021. – 784 с.

8. ENISA. Good Practices For Supply Chain Cybersecurity. – European Union Agency for Cybersecurity, 2023. – 98 p.

9. ГОСТ Р МЭК 62443-3-3-2016. Безопасность систем промышленной автоматизации. Требования к системам и компонентам. – М.: Стандартинформ, 2016.

10. Гайдмака В.А. Кибербезопасность автоматизированных систем управления технологическими процессами. – М.: Горячая линия – Телеком, 2020. – 312 с.

11. Васильев В. И., Кириллова А. Д., Кухарев С. Н. Кибербезопасность автоматизированных систем управления промышленных объектов (современное состояние, тенденции) // Вестник УрФО. Безопасность в информационной сфере. 2018. № 4(30). С. 66–74. DOI: 10.14529/secur180410.

12. Приказ ФСТЭК России от 25 декабря 2017 г. № 239 «Об утверждении требований по обеспечению безопасности значимых объектов критической информационной инфраструктуры Российской Федерации».

13. Федеральный закон от 19 июля 2017 г. № 187 «О безопасности критической информационной инфраструктуры Российской Федерации».

References

1. Bykasov A. V., Sokolov A. N., Barinov A. E. Trust Conditions for Active Scanning Methods for Finding

Vulnerabilities in ICS Networks // 2023 International Russian Automation Conference (RusAutoCon). IEEE, 2023. DOI: 10.1109/RusAutoCon58002.2023.10272827.

2. Bykasov A. V., Sokolov A. N. Statisticheskaya model' otsenki vliyaniya kombinirovannogo metoda aktivnogo skanirovaniya na stabil'nost' funktsionirovaniya seti ASU TP // Vestnik UrFO. Bezopasnost' v informatsionnoy sfere. 2024. № 2(52). S. 68–76. DOI: 10.14529/secur240207.

3. Kleynrok L. Teoriya massovogo obsluzhivaniya. – M.: Mashinostroyeniye, 1979. – 432 s.

4. Bronshteyn O.I., Dukhovnyy I.M. Modeli prioritetnogo obsluzhivaniya v informatsionno-vychislitel'nykh sistemakh. – M.: Nauka, 1976. – 220 s.

5. Gnedenko B.V., Kovalenko I.N. Vvedeniye v teoriyu massovogo obsluzhivaniya. – 2-ye izd. – M.: Nauka, 1987. – 336 s.

6. Bertsekas D., Gallager R. Data Networks. – 2nd ed. – Prentice Hall, 1992. – 640 p.

7. Stollings V. Sovremennyye komp'yuternyye seti. – 4-ye izd. – SPb.: Piter, 2021. – 784 s.

8. ENISA. Good Practices For Supply Chain Cybersecurity. – European Union Agency for Cybersecurity, 2023. – 98 p.

9. GOST R MEK 62443-3-3-2016. Bezopasnost' sistem promyshlennoy avtomatizatsii. Trebovaniya k sistemam i komponentam. – M.: Standartinform, 2016.

10. Gaydmaka V.A. Kiberbezopasnost' avtomatizirovannykh sistem upravleniya tekhnologicheskimi protsessami. – M.: Goryachaya liniya – Telekom, 2020. – 312 s.

11. Vasil'yev V. I., Kirillova A. D., Kukharev S. N. Kiberbezopasnost' avtomatizirovannykh sistem upravleniya promyshlennykh ob"yektov (sovremennoye sostoyaniye, tendentsii) // Vestnik UrFO. Bezopasnost' v informatsionnoy sfere. 2018. № 4(30). S. 66–74. DOI: 10.14529/secur180410.

12. Prikaz FSTEK Rossii ot 25 dekabrya 2017 g. № 239 «Ob utverzhdenii trebovaniy po obespecheniyu bezopasnosti znachimykh ob"yektov kriticheskoy informatsionnoy infrastruktury Rossiyskoy Federatsii».

13. Federal'nyy zakon ot 19 iyulya 2017 g. № 187 «O bezopasnosti kriticheskoy informatsionnoy infrastruktury Rossiyskoy Federatsii».

СОКОЛОВ Александр Николаевич, кандидат технических наук, доцент, заведующий кафедрой «Защита информации» федерального государственного автономного образовательного учреждения высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, Уральский федеральный округ, Челябинская область, г. Челябинск, просп. В.И. Ленина, д. 76. E-mail: sokolovan@susu.ru

БЫКАСОВ Андрей Витальевич, аспирант федерального государственного автономного образовательного учреждения высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, Уральский федеральный округ, Челябинская область, г. Челябинск, просп. В.И. Ленина, д. 76. E-mail: andreybikasov@gmail.com

SOKOLOV Alexander Nikolaevich, Candidate of Technical Sciences, Associate Professor, Head of the Information Security Department of the Federal State Autonomous Educational Institution of Higher Education South Ural State University (National Research University). 454080, Ural Federal District, Chelyabinsk Region, Chelyabinsk, prosp. IN AND. Lenina, d. 76. E-mail: sokolovan@susu.ru

BYKASOV Andrey Vitalievich, post-graduate student of the Federal State Autonomous Educational Institution of Higher Education "South Ural State University (National Research University)". 454080, Ural Federal District, Chelyabinsk Region, Chelyabinsk, prosp. IN AND. Lenina, d. 76. E-mail: andreybikasov@gmail.com

УДК 519.242.7 Вестник УрФО № 2(60) / 2026, с. 32–46

Поршнев С.В., Сафиуллин Н.Т.

DOI: 10.14529/secur260204

АНАЛИЗ ОПЫТА ПРИМЕНЕНИЯ МЕТОДА ФАКТОРНОГО ЭКСПЕРИМЕНТА В ЗАДАЧАХ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Проведен анализ опыта применения метода факторного эксперимента в задачах информационной безопасности, результаты которого показали, что данный метод, несмотря на его потенциальные возможности для тестирования средств защиты и систем защиты информации не используется. Для исправления этой ситуации, по мнению авторов, целесообразно разработать научно обоснованную методику, обеспечивающую отображение семантических описаний угроз безопасности информации, уязвимостей средств и систем в количественные значения входных факторов (варьируемых показателей), а также оценку влияния данных показателей на безопасность защищаемой информационной инфраструктуры.

Ключевые слова: *метод факторного эксперимента, полный факторный эксперимент, дробный факторный эксперимент, планирование факторного эксперимента, информационная безопасность.*

Porshnev S.V., Safiullin N.T.

ANALYSIS OF THE FACTORIAL EXPERIMENT METHOD USAGE IN TASKS OF INFORMATION SECURITY

In this paper an analysis of current usage of the factorial experiment method in information security field is presented. The results show that this method has a potential usage in testing protective means and defensive systems, but currently underused. To improve its usability, it is suggested to develop a scientifically sound methodology that ensures the mapping of semantic descriptions of potential threats in information security and vulnerabilities into quantitative values of input factors (varying parameters), as well as an assessment of the influence of these factors on the safety of the protected information infrastructure.

Keywords: *factorial experiment method, full factorial experiment, fractional factorial experiment, full factorial design, information security.*

1. Введение

В современных условиях задача оценки безопасности[1] информационной инфраструктуры (совокупности объектов информатизации, информационных систем, сайтов в сети Интернет и сетей связи, расположенных на территории РФ, а также на территориях, находящихся под юрисдикцией РФ или используемых на основании международных договоров РФ[2]), а также объектов критической информационной инфраструктуры[3] (КИИ) и различных программных приложений. (Соответственно, понятие «опасность» для объекта можно определить как реальное (существующее, действующее) или потенциально возможное состояние объекта, которое может привести к ущербу в существовании, функционировании, развитии объекта[4]).

Для решения данной задачи, сегодня, наиболее популярными оказываются подходы, основанные на методах теории оценки рисков, методов интеллектуального анализа данных, формальных методах информационной безопасности, статистическом моделировании кибератак (см., например, [1–12]).

Для оценки безопасности информационной инфраструктуры также активно используется методология «черного ящика» (ЧЯ) (англ., *Black Box*) [13], которая наиболее часто, как показали результаты информационного поиска в сети Интернет, используется для динамического тестирования средств защиты информации (СрЗИ), а также программного обеспечения (англ. *Dynamic Application Security Testing, DAST*) с целью обнаружения в них уязвимостей (проведения пентестов), которые может использовать нарушитель для проведения компьютерных атак (КА). Среди инструментов динамического тестирования безопасности, включенных в реестр отечественного программного обеспечения, отметим: *PT Application Inspector, PT BlackBox (Positive Technologies); AppChecker Cloud* (АО «НПО «Эшелон»); *SafeERP* («Газинформсервис»); *BI.ZONE* Анализ защищенности приложений (компания *BI.ZONE*); *Linx SAST* (компания *Linx*); *Solar AppScreener* (компания *Solar*); Айтуби (компания Айтуби).

Напомним, что методология ЧЯ, была разработана для построения статистических моделей сложных систем, формализованное описание структуры которых неизвестно системы (в том числе, неизвестны структура системы, механизмы и правила взаимодействия, входящих в нее подсистем, отсутствует математическая модель системы, разработанная на основе использования аналитических методов, например, вследствие отсутствия соответствующих физических моделей системы и/или адекватного математического аппарата), на основе результатов экспериментальных исследований, называемых факторным экспериментом[5] (ФЭ), проводимых в соответствие с планом факторного эксперимента (см., например, [14], где приведены ссылки на соответствующие монографии, в которых заложены теоретические основы метода ФЭ).

Однако, анализ перечисленных выше инструментов динамического тестирования показывает, что ни один из них не предоставляет пользователю функционала для планирования, проведения и анализа результатов ФЭ, на основе модели ЧЯ. С нашей точки зрения, данная ситуация, в первую очередь, обусловлена отсутствием способов отображения векторов КА, представляющих собой некоторые формальные текстовые описание в виде строк, содержащих соответствующие шифры КА, во входные выходные и факторы ЧЯ, измеряемые в количественных шкалах, а также способов обратного отображения количественных входных и выходных факторов в понятийное пространство ИБ и их интерпретации в данном пространстве.

В этой связи авторы:

Изучили опыт использования ФЭ для решения задач ИБ.

[1] В соответствие с Толковым словарем русского языка: «безопасность» (объекта) – это состояние (объекта), при котором (объекту) не угрожает опасность, имеется защита (объекта) от опасности; «опасность» – возможность, угроза чего-либо очень плохого, какого-либо несчастья; угроза – возможная опасность.

[2] Доктрина информационной безопасности РФ. Утверждена Указом Президента РФ от 5 декабря 2016 г. № 646 «Об утверждении Доктрины информационной безопасности Российской Федерации». Подпункт «з» пункта 2.

[3] Федеральный закон «О безопасности критической информационной инфраструктуры Российской Федерации» от 26.07.2017 № 187-ФЗ (последняя редакция).

[4] В действующих НПА в области ИБ также используется понятие «фактор, воздействующий на защищаемую информацию», смежное с понятием «угроза» (см. например, «Выписку из Концепции государственной системы обнаружения, предупреждения и ликвидации компьютерных атак на информационные ресурсы Российской Федерации», утвержденной Президентом РФ 12 декабря 2014 г. № К-1274).

[5] «Фактор» (лат. factor «делающий, производящий») – причина, движущая сила какого-либо процесса, определяющая его характер или отдельные его черты.//Большая советская энциклопедия : [в 30 т.] / гл. ред. А. М. Прохоров. –3-е изд. –М. : Советская энциклопедия, 1969-1978. [Электронный ресурс]. - URL: https://bigenc.ru/c/eksperiment-509478 (дата обращения: 17.11.2024).

Провели самостоятельный анализ источников информации о КА и технологиях их реализации, в том числе:

– каталог известных уязвимостей ИБ *MITRE*[6] *CVE* (*Common Vulnerabilities and Exposures*) [15];

– базу знаний *MITRE ATT&CK* (*Adversarial Tactics, Techniques & Common Knowledge*) [16], содержащую модели реальных сценариев компьютерных атак (модели поведения злоумышленников), которая собрана на основе анализа тысяч инцидентов ИБ;

– базу знаний *MITRE D3FEND* (де-факто, базу защитных техник) [17], базу данных угроз безопасности информации БДУ ФСТЭК России [18];

– Методический документ «Методика оценки угроз безопасности информации» [19].

Отметим, что в известных работах [20–], посвященных анализу контента информации размещенной в каталоге *MITRE CVE* и базах знаний *MITRE ATT&CK, MITRE D3FEND*, с точки зрения возможности ее использования для планирования ФЭ, ранее не проводился.

Научная новизна настоящей статьи состоит в разработки модели ЧЯ, используемой далее для проведения ФЭ с целью оценки безопасности информационно-телекоммуникационных систем и средств защиты информации, в терминах *MITRE CVE, MITRE ATT&CK, MITRE D3FEND*.

2. Анализ опыта применения ПФЭ в задачах ИБ

Рассмотрим примеры задач ИБ, для решения которых использовался метод ФЭ.

1) Для оценки вероятностей угроз ИБ [26].

Здесь для некоторого анонимного предприятия Z оценивалась вероятность проникновения нарушителя в помещение, в котором находится ИС, в предположении, что на реализацию угрозы оказывают влияние следующие факторы рисков:

1. X_1 – выход из строя электронного замка;
2. X_2 – выход из строя видеокамер;
3. X_3 – инсайдер, который может сообщить нарушителю о выходе из строя электронного замка и/или видеокамер.
4. X_4 – сотрудник, который может случайно пропустить нарушителя в помещение, в котором располагается информационная система.

В связи с тем, что модель ПФЭ в этом случае состояла из 16 опытов, авторы использовали модель дробного факторного эксперимента, состоящего из 8 опытов, матрица которого представлена в таблице 1.

Таблица 1

Таблица плана факторного эксперимента

№	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	X_4 $X_1X_2X_3$	*Y*
1	+1	–1	–1	–1	+1	+1	+1	–1	0
2	+1	+1	–1	–1	–1	–1	+1	+1	0,1
3	+1	–1	+1	–1	–1	+1	–1	+1	0,1
4	+1	+1	+1	–1	+1	–1	–1	–1	0,2
5	+1	–1	–1	+1	+1	–1	–1	+1	0,8
6	+1	+1	–1	+1	–1	+1	–1	–1	0,85
7	+1	–1	+1	+1	–1	–1	–1	+1	0,9
8	+1	+1	+1	+1	+1	+1	+1	+1	1,0

Входным факторам X_n, $n = \overline{1,4}$ присваивались значения ± 1 (+1 – угроза реализуется, –1 – угроза отсутствует). Значения выходного показателя Y_1, $i = \overline{1,8}$ (вероятности угроз получения доступа нарушителя к информационной системе) при сочетаниях входных факторов, указанных в таблице 1, оценивалась экспертом.

Оказалось, что функция отклика описывается следующим выражением:

$Y = 0,5 + 0,0375X_1 + 0,0625X_2 + 0,385X_3 + 0,0125X_4$,

из которого видно, что наибольшее влияние на функцию отклика оказывает фактор X_3, наименьшее фактор – X_4.

2) Для оценки угроз и рисков информационной безопасности на примере оценки вероятности внедрения внутренним нарушите-

[6] MITRE – американская некоммерческая организация, взаимодействующая с оборонными и гражданскими ведомствами США, которая разрабатывает модели, стандарты, инструменты, базы знаний в области кибербезопасности.

лем вредоносного кода [27]. Здесь для некоторого предприятия Z оценивалась вероятность реализации внедрения внутренним нарушителем вредоносного кода в предположении, что на реализацию угрозы оказывают влияние следующие факторы рисков:

1. X_1 – устаревшие антивирусные базы;

2. X_2 – отсутствие запрета на запуск исполняемых файлов от имени пользователей;

3. X_3 – возможность управления функционированием антивирусного ПО от имени пользователей;

4. X_4 – возможность установления удаленного подключения к персональному компьютеру.

Для построения модели, описывающей зависимость вероятности внедрения внутренним нарушителем вредоносного кода Y от выбранных факторов X_1, X_2, X_3, X_4. В связи с тем, что модель ПФЭ в этом случае состояла из 16 опытов, авторы использовали модель дробного факторного эксперимента, состоящего из 8 опытов, план которого представлен в таблице 2.

Таблица 2

Таблица плана ПФЭ

№	X_0	X_1	X_2	X_3	X_4 $X_1X_2X_3$	X_1X_2	X_1X_3	X_2X_3	Y
1	+1	−1	−1	−1	−1	+1	+1	+1	0
2	+1	+1	−1	−1	+1	−1	−1	+1	0,95
3	+1	−1	+1	−1	+1	−1	+1	−1	0,87
4	+1	+1	+1	−1	−1	+1	−1	−1	0,97
5	+1	−1	−1	+1	+1	+1	−1	−1	0,82
6	+1	+1	−1	+1	−1	−1	+1	−1	0,96
7	+1	−1	+1	+1	−1	−1	−1	+1	1,0
8	+1	+1	+1	+1	+1	+1	+1	+1	0,8

Входным факторам X_n, $n=\overline{1,4}$ присваивались значения ± 1 (+1 – угроза реализуется, −1 – угроза отсутствует). Значения выходного показателя Y_i, $i=\overline{1,8}$ (вероятности угроз получения доступа нарушителя к информационной системе) при сочетаниях входных факторов, указанных в таблице 2, оценивалась экспертом.

Оказалось, что функция отклика описывается следующим выражением:

$$Y = 0.82 + 0,15X_1 + 0,14X_2 + 0,12X_3 + 0,09X_4 - 0,13X_1X_2 - \\ - 0,12X_1X_3 - 0,9X_2X_3,$$

из которого видно, что наибольшее влияние на функцию отклика оказывают факторы X_1, X_2, наименьшее – фактор X_4.

3) Для оценки факторов, влияющих на реализацию угроз и рисков информационной безопасности [28]. Здесь была использована модель угроз безопасности информации, обрабатываемой в системе электронного документооборота, составленная в соответствии с требованиями Методического документа «Методика оценки угроз безопасности информации» (утв. ФСТЭК от 5 февраля 2021 г.) [19] и на основании банка угроз [18]. Оценку рисков выполняли для группы угроз, реализуемых в государственной организации внутренним нарушителем с низким потенциалом, что приводит к нарушению конфиденциальности, целостности и доступности информации [18]:

– УБИ.074 – угроза несанкционированного доступа к аутентификационной информации;

– УБИ.086 – угроза несанкционированного изменения аутентификационной информации;

– УБИ.088 – угроза несанкционированного копирования защищаемой информации;

– УБИ.152 – угроза удаления аутентификационной информации;

– УБИ.167 – угроза заражения компьютера при посещении неблагонадёжных сайтов;

–УБИ.168 – угроза «кражи» учётной записи доступа к сетевым сервисам;

– УБИ.172 – угроза распространения «почтовых червей»;

– УБИ.191 – угроза внедрения вредоносного кода в дистрибутив программного обеспечения.

В качестве причин реализации угрозы несанкционированного доступа к аутентификационной информации (УБИ.074 [16]) были рассмотрены следующие факторы:

1) X_1 – небрежность действий персонала в информационном измерении;

2) X_2 – использование устаревших антивирусных баз;

3) X_3 – отсутствие запрета на запуск исполняемых файлов от имени пользователей;

4) X_4 – возможность управления функционированием антивирусного программного обеспечения от имени пользователей.

Матрица плана проведенного 2^3-дробного факторного эксперимента представлена в таблице 3.

Вероятность реализации угроз для каждого сценария $\left(Y_{i,j}\right)$ оценивалась экспертным методом (метод непосредственного оценивания) по шкале от 0 до 1. Объектам экспертизы (каждому сценарию) каждым экспертом присваивались баллы, при этом наиболее значимому объекту (сценарию) присваивалось наибольшее количество баллов в соответствие с выбранной шкалой оценок. Вероятность реализации угроз для каждого сценария $\left(Y_i\right)$ оценивалась как среднее арифметическое значение оценок сценариев каждым из четырех экспертов. Результаты оценки вероятностей реализации угроз методом экспертных оценок представлены в табл. 4.

На основании результатов оценки вероятностей реализации угрозы (табл. 4) для восьми планов проведения 4-х факторных

Таблица 3

Таблица плана факторного эксперимента

№	X_0	X_1	X_2	X_3	X_4	X_1X_2	X_1X_3	X_2X_3	Y
1	+1	−1	−1	−1	−1	+1	+1	+1	Y_1
2	+1	+1	−1	−1	+1	−1	−1	+1	Y_2
3	+1	−1	+1	−1	+1	−1	+1	−1	Y_3
4	+1	+1	+1	−1	−1	+1	−1	−1	Y_4
5	+1	−1	−1	+1	+1	+1	−1	−1	Y_5
6	+1	+1	−1	+1	−1	−1	+1	−1	Y_6
7	+1	−1	+1	+1	−1	−1	−1	+1	Y_7
8	+1	+1	+1	+1	+1	+1	+1	+1	Y_8

Таблица 4

Результаты оценки вероятностей реализации угроз методом экспертных оценок

№	Ранги $R_{i,j}$, присвоенные экспертами				$\sum_{i=1}^{2^3} R_{i,j}$	Y_j
1	0,30	0,30	0,30	0,40	1,30	0,33
2	0,50	0,50	0,50	0,60	2,10	0,53
3	0,40	0,60	0,50	0,50	2,00	0,50
4	0,50	0,70	0,60	0,70	2,50	0,63
5	0,80	0,70	0,80	0,60	2,90	0,73
6	0,70	0,80	0,80	0,80	3,10	0,78
7	0,85	1,00	0,90	0,70	3,45	0,86
8	1,00	1,00	0,95	1,00	3,95	0,99

экспериментов, представленных в табл. 3, получена следующая функции отклика:

$$Y = 0,669 + 0,064X_1 + 0,076X_2 + 0,171X_3 + 0,001X_1X_2 - 0,019X_1X_3 + 0,009X_2X_3.$$

Авторы интерпретировали найденное выражение для функции отклика следующим образом: значение $b_0 = 0,669 > 0,5$ свидетельствует об актуальности рассматриваемой угрозы и значимости влияния каждого из четырех рассмотренных факторов на реализацию угрозы информационной безопасности. Наибольший вклад в реализацию угрозы информационной безопасности вносит влияние фактор X_3 – отсутствие запрета на запуск исполняемых файлов от имени пользователей, который также усиливает и усиливает негативное действие фактора X_2 – использование устаревших антивирусных баз, а также уменьшает влияние X_1 – (небрежность действий персонала в информационном измерении). Наименьшее влияние на вероятность реализации угрозы фактор X_4 – возможность управления функционированием антивирусного программного обеспечения от имени пользователей.

4) Для разработки методики оценки эффективности системы менеджмента информационной безопасности (СМИБ) по времени реакции системы на инциденты ИБ [28]. В качестве интегрального показателя эффективности системы менеджмента информационной безопасности предложено использовать время реакции СМИБ на инциденты ИБ, вычисляемое по формуле:

$$T_{\text{рсИБ}} = \sum_{i=1}^{k} \left(T_{\text{обнИБ}_i} + T_{\text{блокИБ}_i} + T_{\text{регИБ}_i} + T_{\text{оповИБ}_i} \right),$$

где k – количество атрибутов, подлежащих измерению; $T_{\text{рсИБ}}$ – время реакции СМИБ на инцидент ИБ; $T_{\text{обнИБ}_i}$ – время обнаружения инцидента ИБ; $T_{\text{блокИБ}_i}$ – время блокирования несанкционированных действий по нарушению ИБ; $T_{\text{регИБ}_i}$ – время регистрации события, связанного с инцидентом ИБ; $T_{\text{оповИБ}_i}$ – время оповещения персонала службы ИБ о выявленном инциденте ИБ.

Для расчета показателя автор рекомендовал разделить меры и средства управления и контроля для выполнения проверок на группы по времени реагирования на инцидент ИБ, и использовать следующие виды группирования:

– немедленный контроль (время реакции – несколько минут ($T_{\text{рсИБ}} \le 10$ мин));

– суточный контроль (время реакции – в течение суток ($T_{\text{рсИБ}} \le 24$ ч));

– контроль изменения политик ИБ (время реакции – в течение нескольких суток ($T_{\text{рсИБ}} \le 10$ сут)).

В связи с тем, что для исследованных мер и средств управления и контроля СМИБ функция отклика (аналитической модели), описывающая показатель эффективности СМИБ, получить ее аналитического описания не удалось, автор, предложил использовать метод ПФЭ или ДФЭ, однако, не привел ни каких-либо наборов варьируемых факторов, ни диапазонов их изменений.

Предложенная в [28] методика расчета показателя эффективности СМИБ реализуется выполнением следующей последовательности действий.

1. Выбор необходимого количества атрибутов k и требуемого числа измерений N.

2. Выбор, исходя из априорных данных и экспертных оценок, модели, описывающей, описывающей функциональную зависимость времени реакции СМИБ на инциденты ИБ $T_{\text{рсИБ}}$.

3. Определение интервалов варьирования входных факторов с учетом особенностей СМИБ организации.

4. Выбор ПФЭ (для случая $k \le 3$) или ДФЭ (для случая $k > 3$).

5. Составление матрицы плана ФЭ, удовлетворяющей требованиям D-оптимального плана, и определение, в соответствие с выбранным планом, правила вычисления коэффициентов регрессии b_j.

6. Проведение измерения значений выходного показателя для значений входных факторов, указанных в матрице плана ФЭ.

7. Расчет по результатам измерений коэффициентов регрессии b_j.

8. Проверка статистической значимости коэффициентов регрессии b_j, а также адекватности построенной математической модели.

При этом, однако, автор не привел ни одного конкретного примера использования предложенной методики.

Таким образом, результаты анализа опыта использования метода факторного эксперимента в ИБ позволяют сделать следующие выводы.

1. Метод факторного эксперимента использовался ранее для:

1) оценки вероятности проникновения нарушителя в помещение, в котором находится информационная система;

2) вероятностей угроз и рисков ИБ на примере оценки вероятности внедрения

внутренним нарушителем вредоносного кода;

3) ранжирования факторов, влияющих на реализацию угроз ИБ;

4) разработки методики оценки эффективности СМИБ.

2. В качестве входных варьируемых факторов в данных экспериментах использовались следующие группы факторов:

1) выход из строя электронного замка (X_1); выход из строя видеокамер (X_2); инсайдер, который может сообщить нарушителю о выходе из строя электронного замка и/или видеокамеру (X_3); сотрудник, который может случайно пропустить нарушителя в помещение, в котором располагается информационная система (X_4);

2) устаревшие антивирусные базы (X_1); отсутствие запрета на запуск исполняемых файлов от имени пользователей (X_2); возможность управления функционированием антивирусного ПО от имени пользователей (X_3); возможность установления удаленного подключения к персональному компьютеру (X_4);

3) небрежность действий персонала в информационном измерении (X_1); использование устаревших антивирусных баз (X_2); отсутствие запрета на запуск исполняемых файлов от имени пользователей (X_3); возможность управления функционированием антивирусного программного обеспечения от имени пользователей (X_4);

4) входные варьируемые факторы не указаны.

3. Значения входных варьируемых факторов измерялись по дихотомической шкале (+1 – данная угроза безопасности ИБ присутствует, –1 – данная угроза безопасности информации отсутствует).

4. Значения выходного показателя, характеризующего состояния ИБ изучаемой системы, определялись на основе экспертных оценок, но не результатов их измерений реальных факторных экспериментах.

5. Публикаций с обсуждением результатов применения методов ПФЭ и ДФЭ для тестирования средств защиты и систем защиты информации обнаружить не удалось.

3. Модель «черного ящика» в терминах матриц MITRE ATT&CK и MITRE D3FEND

Предваряя обсуждение модели ЧЯ в терминах матриц MITRE ATT&CK и MITRE D3FEND, рассмотрим классическую постановку задачи, для решения которой используется модель ЧЯ. Имеется система, рассматриваемая как некоторое изолированное целое, содержащее совокупность процессов и средств их реализации, представляющая собой ЧЯ, на который воздействует некоторое конечное множество независящих друг от друга контролируемых (задаваемых) воздействий $\{x_n\}$, $n=\overline{1,N}$ – входных факторов, а также некоторое конечное множество контролируемых неизменяемых факторов (условия, среда функционирования и т.п.) $\{c_k\}$, $k=\overline{1,K}$ и некоторое конечное множество неконтролируемых или контролируемых, но неуправляемых факторов (случайные или неидентифицируемые воздействия, изменения, «дрейф» параметров и т.п.) $\{u_l\}$, $l=\overline{1,L}$. При этом состояние ЧЯ, как полагают, априори, можно описать некоторым конечным множеством, измеряемых независимых друг от друга показателей, $\{Y_m\}$, $m=\overline{1,M}$, называемых откликами (рис. 1).

В наиболее общей математической форме сформулированные выше исходные предположения (аксиомы) относительно факторов $\overline{X}$, $\overline{C}$ и $\overline{U}$, определяющих значения откликов «черного ящика» $\overline{Y}$, записываются в следующем виде:

$$\overline{Y}=\overline{F}\left(\overline{X},\overline{C},\overline{U}\right), \qquad (1)$$

где: $\overline{Y}=(y_1,y_2,\ldots,y_M)$ – вектор размерности M, содержащий значения выходных оцениваемых (измеряемых) показателей, характеризующих состояние «черного ящика, которые выбираются, исходя из целевых аспектов проводимого экспериментального исследования;

$\overline{X}=(x_1,x_2,\ldots,x_N)$ – вектор размерности N, содержащий значения варьируемых факторов;

$\overline{C}=(c_1,c_2,\ldots,c_K)$ – вектор размерности K, содержащий значения контролируемых, но неизменяемых факторов исследования;

$\overline{U}=(u_1,u_2,\ldots,u_L)$ – вектор размерности L, содержащий значения неконтролируемых или контролируемых, но не управляемых (случайных, неидентифицируемых) факторов;

$\overline{F}$ – в общем виде некоторая векторфункция (в наиболее общем случае, функционал), ставящая в соответствие векторам $\overline{X},\overline{C},\overline{U}$, вектор $\overline{Y}$, содержащий значения (выходных) показателей состояния «черного ящика».

Говорят, что множество векторов $\overline{F}$ образует пространство выходных показателей.

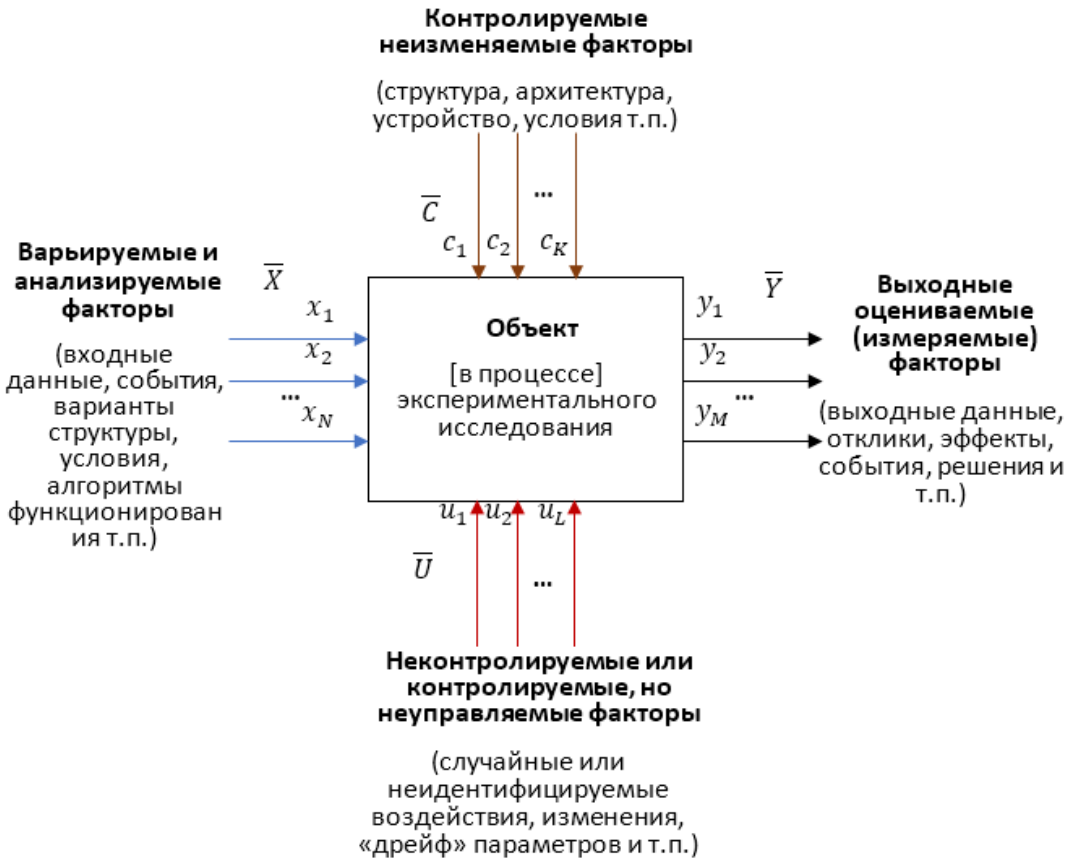

Рис. 1. Модель объекта в процессе экспериментального исследования в идеологии «черного ящика

Отметим главное требование к варьируемым и контролируемым факторам – они должны быть принципиально наблюдаемыми (измеряемыми, фиксируемыми), значимыми и существенными (оказывающими влияние на состояние объекта исследования), а также независимыми друг от друга (эквивалентно, некоррелированными друг с другом). При этом природа факторов и соответственно шкалы их измерения могут быть совершенно различными: количественными, качественными, дихотомическими, структурно-образными и т.д. (Например, при тестировании СЗИ могут определяться или вычисляться вероятности обнаружения атак, т.е. количественные факторы; при исследованиях защищенности – наличие и идентификация уязвимостей, т.е. дихотомически-назывные факторы, и т.д.). В этой связи на первом этапе планирования экспериментов все факторы и выходные показатели должны быть отображены в соответствующие цифровые значения. Относительно неконтролируемых факторов принимается постулат о том, что в ходе проводимых экспериментов их значения остаются постоянными.

Цель проведения экспериментального исследования состоит в идентификации неизвестной функции (функционала) $\overline{F}$ в (1) на основе использования задаваемых количественных значений входных факторов $\left(\overline{X}\right)$, известных количественных значений контролируемых, но неизменяемых факторов $\left(\overline{C}\right)$, неконтролируемых или контролируемых, но не управляемых факторов $\left(\overline{U}\right)$ и измеренных количественных значений отклика «черного ящика» $\overline{Y}$.

Таким образом, что основная проблема подготовки ФЭ состоит в выборе входных факторов (внешних воздействий) и выходных показателей, характеризующих состояние защищаемой информационной системы, с целью отображения решаемой задачи ИБ в пространство «классической» модели ЧЯ. Для ее решения был проведен анализ известных разработок в области ИБ, реализованных организацией MITRE, в том числе: каталог известных уязвимостей ИБ CVE [15], базу знаний ATT&CK [16], содержащую модели реальных сценариев компьютерных атак (модели поведения злоумышленников), которая собрана на основе анализа тысяч инцидентов ИБ, и

базу знаний D3FEND (да-факто, базу защитных техник, используемых для отражения КА) [17]. Отметим, что необходимость сопоставления моделей противодействия КА и конкретных мер защиты, которые должен применять специалист по ИБ, была осознана в процессе создания базы знаний MITRE D3FEND, о чем впервые было заявлено в [30]. Здесь авторы привели таблицу, в которой были указаны ключевые аспекты сопоставления моделей противодействия КА и мер защиты, представленную на рис. 2, из которого видно, что ключевые аспекты сопоставления моделей противодействия КА и мер защиты, опирающихся на семантику IDEF-моделей, аналогичны элементам модели «черного ящика», представленной на рис. 1.3.1, используемой в ФЭ.

Действительно, входные воздействия в

TABLE I
TECHNIQUE ELICITATION QUESTIONS

How does the technology work?	Activity model element
What are the data inputs?	Activity Input
What are the data outputs?	Activity Output
When does the analysis occur?	Activity Control
What are the analytical algorithms?	Activity Mechanism
How does it work at scale?	Activity Mechanism
How can it be circumvented?	Activity Mechanism

Рис. 2. Таблица вопросов по определению используемых методик

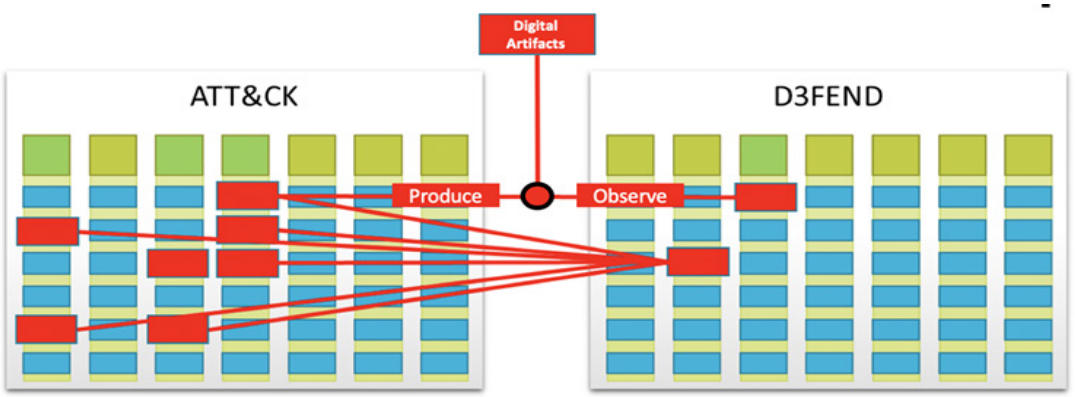

Рис. 3. Соответствие баз знаний MITRE ATT&CK и MITRE D3FEND через концепцию цифровых артефактов

модели «черного ящика» соответствуют Activity Input, выходные данные – Activity Output, контролируемые условия – Activity Control, механизмы/алгоритмы – Activity Mechanism. Однако, решение обсуждаемой задачи, как и задачи (1) (см. рис. 1) «в лоб» оказывается невозможным из-за их высоких размерностей и, как следствие, высоких временных затрат, поэтому обсуждаемая модель была декомпозирована в виде онтологии цифровых артефактов.

Для концептуального моделирования защитных техник в базу знаний MITRE D3FEND была введен новая ключевая конструкция, названная цифровым **артефактом** (Digital Artifact), а в систему их представления – онтология D3FEND DAO. В соответствие с [30], цифровой артефакт – это такой цифровой объект системы, с которым взаимодействуют атакующая систему и защищающая систем стороны, изменяя его состояние. При этом атакующая сторона создает (применяет) факторы (в терминах факторного эксперимента – варьируемые показатели) с целью изменения состояния данного цифрового артефакта (в терминах факторного эксперимента – выходные показатели), а защищающая сторона активно наблюдает за состоянием цифрового артефакта (в терминах факторного эксперимента – измеряет значения его выходных показателей) и реагирует на внешние факторы с целью обеспечения защиты (неизменяемости) цифрового артефакта или возвращения его состояния в исходное. Описанный механизм обеспечил однозначное соответствие между сценариями КА, находящихся в базе знаний MITRE ATT&CK (вход цифрового артефакта), и методик защит, находящихся в базе

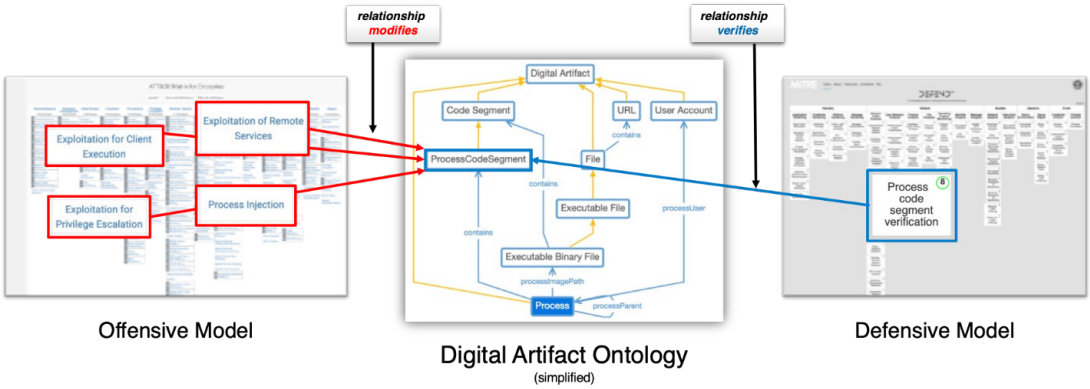

Рис. 4. Связь матриц знаний MITRE ATT&CK и MITRE D3FEND через онтологию цифровых артефактов D3FEND DAO [31]

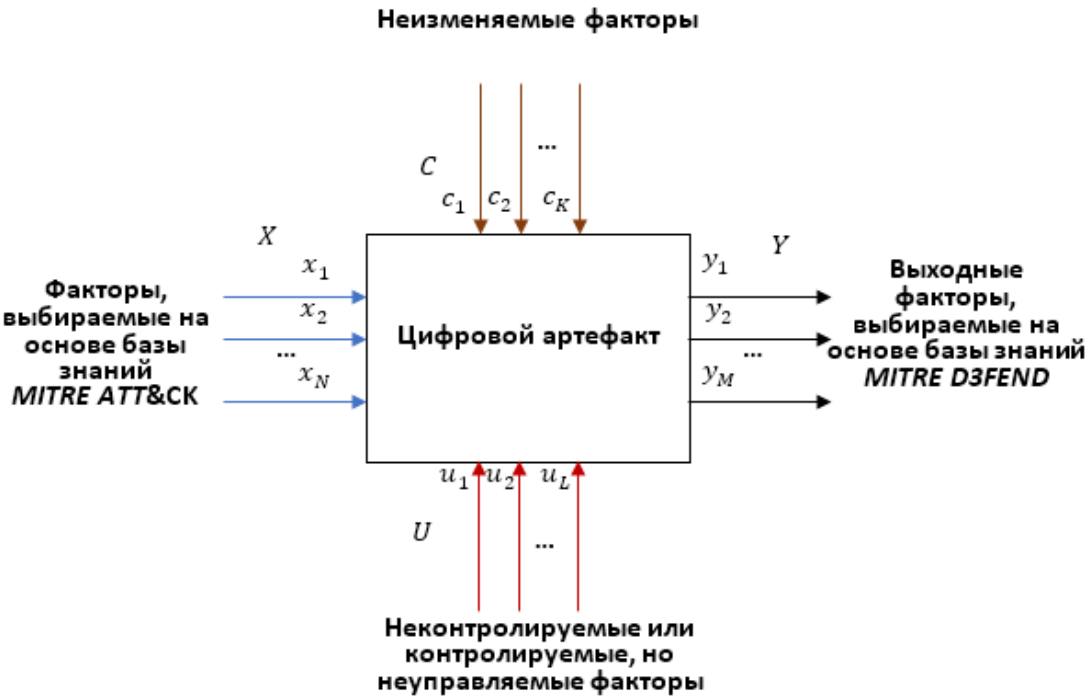

Рис. 5. Модель ЧЯ в терминах матриц MITRE ATT&CK и MITRE D3FEND

знаний MITRE D3FEND (выход цифрового артефакта) (см. рис. 3).

Отметим, что в исходной версии онтологии D3FEND DAO, артефакт в большей степени рассматривался как некоторый цифровой след, который анализировали специалисты в области ИБ, то есть, де-факто, он использовался для расследования инцидентов ИБ, но не для решения задачи противодействия атакам на ранних этапах защиты системы, либо при разработке средств защиты информации. В этой связи далее входы цифровых артефактов были связаны с базой знаний MITRE ATT&CK не только через эффект «создания» или «реализации», но и через любые другие способы воздействия на них, аналогично, независимым факторам в модели ЧЯ. При этом, принцип исключительного «наблюдения» факторов на выходе цифровых артефактов был заменен принципом «множества отношений», привязанных к защитным техникам или средствам защиты (см. пример на рис. 3).

Принимая во внимание приведенное выше определение цифрового артефакта, а также рис. 1, 3, можно предложить следующий прототип модели «черного ящика» в терминах матриц MITRE ATT&CK и MITRE D3FEND (рис. 3).

Однако для практического использования предложенного прототипа модели «черного ящика» в терминах базы знаний MITRE ATT&CK и MITRE D3FEND требуется, в свою очередь, решить следующие задачи.

1) Задачу отображения вербального опи-

сания факторов, используемых в базах знаний MITRE ATT&CK и MITRE D3FEND, в измеряемые количественные значения входных и выходных факторов, наличие которых является необходимым условием построения ПФЭ и его реализации на практике.

2) Задачу, связанную с отсутствием в онтологии цифровых артефактов D3FEND DAO определения понятия «неизменяемые факторы». (По определению цифрового артефакта, для того чтобы объект представлял собой артефакт, нужно иметь возможность наблюдения изменений его состояния, однако, любые неизменяемые факторы находятся за рамками этого термина.

3) Задачу, связанную с отсутствием в базе знаний MITRE D3FEND, как таковых, отдельных неконтролируемых (или неуправляемых) факторов, поскольку входные факторы в момент проведения КА (но не ее тестирования или моделирования) становятся неуправляемыми, так как они зависят от атакующей стороны, но не зависят от стороны защиты. (Таким образом, следует считать, что-либо известны все входные факторы цифрового артефакта, либо существуют неизвестные неконтролируемые факторы, не включенные в базу знаний MITRE D3FEND.)

Обсуждение подходов, обеспечивающих решение перечисленных выше задач, является предметом последующих публикаций.

Заключение

Результаты анализа опыта использования метода ФЭ в ИБ позволяют сделать следующие выводы.

Метод ФЭ использовался для:

1) оценки вероятности проникновения нарушителя в помещение, в котором находится информационная система;

2) вероятностей угроз и рисков ИБ на примере оценки вероятности внедрения внутренним нарушителем вредоносного кода;

3) ранжирования факторов, влияющих на реализацию угроз ИБ;

4) разработки методики оценки эффективности СМИБ.

Публикаций с обсуждением результатов применения методов ПФЭ и ДФЭ для тестирования средств защиты и систем безопасности информационной инфраструктуры обнаружить не удалось.

Обосновано, что основная проблема, препятствующая широкому использованию метода ФЭ для тестирования средств защиты и систем безопасности информационной инфраструктуры, состоит в отсутствии научно обоснованных методик отображения семантических описаний угроз безопасности информации, уязвимостей средств и систем в перечни входных варьируемых факторов, значения которых должны измеряться в количественных шкалах, а также выбора количественных показателей, характеризующих состояние ИБ защищаемой информационной системы.

На первом этапе решения данной проблемы проведен анализ известных разработок в области ИБ, реализованных организацией MITRE, в том числе: MITRE CVE, MITRE ATT&CK, D3FEND, построена модель ЧЯ в терминах матриц MITRE ATT&CK и MITRE D3FEND и определены задачи, требующих решения на последующих этапах исследования.

Литература

1. Файзулаев Д.Ф., Морозов Б.Б. Методы и средства анализа рисков информационной безопасности предприятия// Д.Ф. Файзулаев, Б.Б. Морозов/ Безопасность информационных технологий, 2017. № 3. –С. 69–73. DOI: http://dx.doi.org/10.26583/bit.2017.3.09

2. Васильев В.И., Вульфин А.М., Гузаиров М.Б. Оценка рисков информационной безопасности с использованием нечетких продукционных когнитивных карт//В.И. Васильев, А.М. Вульфин, М.Б. Гузаиров/ Информационные технологии, 2018. Т. 24. Вып. 4 –С. 266–273. DOI: 10.17587/it.24.266-273

3. Васильев В.И., и др. Оценка рисков кибербезопасности АСУ ТП промышленных объектов, на основе вложенных нечетких когнитивных карт//В.И. Васильев, А.М. Вульфин, М.Б. Гузаиров, В.М. Картак, Л.Р. Черняховская/ Информационные технологии, 2020. Т.26. Вып. 4. –С. 213–221. DOI: 10.17587/it.26.213-221

4. Вульфин А.М. Модели и методы комплексной оценки рисков безопасности объектов критической информационной инфраструктуры на основе интеллектуального анализа данных// А.М. Вульфин/Системная инженерия и информационные технологии, 2023. Т. 5, № 4 (13). С. 50–76. DOI: 10.54708/2658-5014-SIIT-2023-no3-p50

5. Васильев В.И., Кириллова А.Д., Кухарев С.Н. Кибербезопасность автоматизированных систем управления промышленных объектов (современное состояние, тенденции)//В.И. Васильев, А.Д. Ки-

риллова, С.Н. Кухарев /Вестник УрФО. Безопасность в информационной сфере, 2018. Т. 30. Вып. 4. –С. 66–74. DOI: 10.14529/secur180410

6. Васильев В.И., Кириллова А.Д., Вульфин А.М. Когнитивное моделирование вектора кибератак на основе меташаблонов CAPEC//В.И. Васильев, А.Д. Кириллова, А.М. Вульфин/ Вопросы кибербезопасности, 2021. Т. 42. Вып. 2. С. 2–16. DOI: 10.21681/2311-3456-2021-2-2-16

7. Васильев В.И., Вульфин А.М., Гузаиров М.Б., Кириллова А.Д. Интервальное оценивание информационных рисков с помощью нечетких серых когнитивных карт//В.И. Васильев, А.М. Вульфин, М.Б. Гузаиров, А.Д. Кириллова/ Информационные технологии, 2018. Т. 24. Вып. 10. –С. 657-664. DOI: 10.17587/it.24.657-664

8. Аникин И.В. Методы и алгоритмы количественной оценки и управления рисками безопасности в корпоративных информационных сетях на основе нечеткой логики//И.В. Аникин/ Системная инженерия и информационные технологии, 2023. Т. 5, Вып. 3. № 12. –С. 93–113. DOI 10.54708/2658-5014-SIIT-2023-no3-p93

9. Макаренко С.И. Критерии и показатели оценки качества тестирования на проникновение//С.И. Макаренко/Вопросы кибербезопасности, 2021. № 3(43). –С. 43–57. DOI:10.21681/2311-3456-2021-3-43-57

10. Смирнов С.И., Еремеев М.А., Магомедов Ш.Г., Изергин Д.А. Критерии и показатели оценивания качества проведения расследования инцидента информационной безопасности при целевой кибератаке// С.И. Смирнов, М.А. Еремеев, Ш.Г. Магомедов, Д.А. Изергин /Russian Technological Journal, 2024. Т. 12. Вып. 3. –С. 25–36. DOI: https://doi.org/10.32362/2500-316X-2024-12-3-25-36

11. Макаренко С. И., Смирнов Г. Е. Методика обоснования тестовых информационно-технических воздействий, обеспечивающих рациональную полноту аудита защищенности объекта критической информационной инфраструктуры//С. И. Макаренко, Г. Е. Смирнов/ Вопросы кибербезопасности. 2021. № 6. (46). –С. 12–25. DOI:10.21681/2311-3456-2021-6-12-25

12. Котенко И.В. и др. Технологии больших данных для корреляции событий безопасности на основе учета типов связей//И.В. Котенко, А.В. Федорченко, И.Б. Саенко, А.Г. Кушнеревич А.Г./ Вопросы кибербезопасности,Ю 2017. №5 (24). –С. 2–16. DOI: 10.21681/2311-3456-2017-5-2-16

13. Организация системы сетевого мониторинга и оценки состояния информационной безопасности объекта / А.Л. Марухленко, К.Д. Селезнёв, М.О. Таныгин, Л.О. Марухленко // Известия Юго-Западного государственного университета, 2019. Т. 23. № 1. –С. 118–129. DOI: 10.21869/2223-1560-2019-23-1118-129.

14. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. –2-е изд., испр. –М.: ФИЗМАТЛИТ, 2012. –816 с.

15. URL: https://CVEform.mitre.org (Дата обращения: 22.12.2025).

16. URL: https://attack.mitre.org/ (Дата обращения: 22.12.2025).

17. URL: https://d3fend.mitre.org/ (Дата обращения: 22.12.2025.

18. Банк данных угроз безопасности информации// URL: https://bdu.fstec.ru/threat (Дата обращения: 22.12.2025).

19. Методический документ «Методика оценки угроз безопасности информации» (утв. ФСТЭК 05.02.2021 г.)// URL: https://www.consultant.ru/document/cons_doc_LAW_378330/ (Дата обращения: 22.12.2025).

20. Kocaman Yu., Gönen S., Barişkan M.A., Karacayilmaz G., Yilmaz E.N. A Novel approach to continuous CVE analysis on enterprise operating systems for vulnerability assessment. International journal of information technology (Singapore), 2022. Vol. 14. № 3. P. 1433–1443.

21. Aghaei E, Al-Shaer E, Shadid W, Niu X. Automated cve analysis for threat prioritization and impact prediction. 2023. arXiv preprint arXiv:2309.03040. DOI: https://doi.org/10.48550/arXiv.2309.03040

22. Manjunatha A, Kota K, Babu AS, et al. Cve severity prediction from vulnerability description-a deep learning approach. Procedia Comput Sci, 2024; 235:3105–17. DOI: 10.1016/j.procs.2024.04.294

23. Sonmeza F. O., C. Hankina, Malacariab P. Dynamics: An Automatic Attack Graph Generation Framework Based on System Topology, CAPEC, CWE, and CVE Databases. Preprint submitted to Computers and Security September 26, 2022. 22 p. URL: https://www.researchgate.net/publication/364096733_Attack_Dynamics_An_Automatic_Attack_Graph_Generation_Framework_Based_on_System_Topology_CAPEC_CWE_and_CVE_Database

24. Тулинов И.С., Губсков Ю.А., Громов Ю.Ю., Шатских В.В. Формирование базы данных уязвимостей программного обеспечения для оперативной нейтрализации угроз// И.С. Тулинов, Ю.А. Губсков, Ю.Ю. Громов Ю.Ю., В.В. Шатских/ Промышленные АСУ и контроллеры, 2018, № 3. – С. 26–30.

25. Paiva R. Cernambi Virgem Ecológico (CVE), un nuevo caucho de extracción natural con calidad y

directo de la selva amazónica a la industria del calzadohttps. DOI://doi.org/10.4067/S0718-07642021000600101

26. Белкин С.А., Белов В.М., Пивкин Е.Н. Применение факторного планирования эксперимента для оценки вероятностей угроз информационной безопасности// Ползуновский вестник, 2014. № 2. – С. 232–234.

27. Плетнев П. В., Белов В. М., Зубков Е. В., Крыжановская О. А. К вопросу об определении угроз и рисков информационной безопасности с использованием сценарного подхода, и факторного планирования эксперимента// Вестник СибГУТИ. 2016. № 4. – С. 12–18.

28. Маркелова Е.Б., Троеглазова А.В. Оценка факторов, оказывающих влияние на реализацию угроз информационной безопасности// ИНТЕРЭКСПО ГЕО-СИБИРЬ, 2023. Т. 6. № 1. – С. 165–170.

29. Шаго Ф.Н. Методика оценки эффективности системы менеджмента информационной безопасности (СМИБ) по времени реакции системы на инциденты ИБ//Научно-технический вестник информационных технологий, механики и оптики (Scientific and Technical Journal of Information Technologies, Mechanics and Optics), 2014. № 4 (92). – С. 115–123.

30. URL://www.anti-malware.ru/files/27-04-2023/28_s.jpg (Дата обращения: 22.12.2025).

31. URL: https://d3fend.mitre.org/technique/d3f:HostReboot/ (Дата обращения: 21.12.2025).

References

1. Fajzulaev D.F., Morozov B.B. Metody i sredstva analiza riskov informacionnoj bezopasnosti predpriyatiya// D.F. Fajzulaev, B.B. Morozov/ Bezopasnost' informacionnyh tekhnologij, 2017. № 3. –S. 67–73. DOI: http://dx.doi.org/10.26583/bit.2017.3.09

2. Vasil'ev V.I., Vul'fin A.M., Guzairov M.B. Ocenka riskov informacionnoj bezopasnosti s ispol'zovaniem nechetkih produkcionnyh kognitivnyh kart//V.I. Vasil'ev, A.M. Vul'fin, M.B. Guzairov/ Informacionnye tekhnologii, 2018. T. 24. Vyp. 4 –S. 266–273. DOI: 10.17587/it.24.266-273

3. Vasil'ev V.I., i dr. Ocenka riskov kiberbezopasnosti ASU TP promyshlennyh ob"ektov na osnove vlozhennyh nechetkih kognitivnyh kart//V.I. Vasil'ev, A.M. Vul'fin, M.B. Guzairov, V.M. Kartak, L.R. CHernyahovskaya/ Informacionnye tekhnologii, 2020. T.26. Vyp. 4. –S. 213–221. DOI: 10.17587/it.26.213–221

4. Vul'fin A.M. Modeli i metody kompleksnoj ocenki riskov bezopasnosti ob"ektov kriticheskoj informacionnoj infrastruktury na osnove intellektual'nogo analiza dannyh// A.M. Vul'fin/Sistemnaya inzheneriya i informacionnye tekhnologii, 2023. T. 5, № 4 (13). –S. 50–76. DOI: 10.54708/2658-5014-SIIT-2023-no3-p50

5. Vasil'ev V.I., Kirillova A.D., Kuharev S.N. Kiberbezopasnost' avtomatizirovannyh sistem upravleniya promyshlennyh ob"ektov (sovremennoe sostoyanie, tendencii)//V.I. Vasil'ev, A.D. Kirillova, S.N. Kuharev / Vestnik UrFO. Bezopasnost' v informacionnoj sfere, 2018. T. 30. Vyp. 4. –S. 66–74. DOI: 10.14529/secur180410

6. Vasil'ev V.I., Kirillova A.D., Vul'fin A.M. Kognitivnoe modelirovanie vektora kiberatak na osnove metashablonov CAPEC//V.I. Vasil'ev, A.D. Kirillova, A.M. Vul'fin/ Voprosy kiberbezopasnosti, 2021. T. 42. Vyp. 2. S. 2–16. DOI: 10.21681/2311-3456-2021-2-2-16

7. Vasil'ev V.I., Vul'fin A.M., Guzairov M.B., Kirillova A.D. Interval'noe ocenivanie informacionnyh riskov s pomoshch'yu nechetkih seryh kognitivnyh kart//V.I. Vasil'ev, A.M. Vul'fin, M.B. Guzairov, A.D. Kirillova/ Informacionnye tekhnologii, 2018. T. 24. Vyp. 10. –S. 657–664. DOI: 10.17587/it.24.657-664

8. Anikin I.V. Metody i algoritmy kolichestvennoj ocenki i upravleniya riskami bezopasnosti v korporativnyh informacionnyh setyah na osnove nechetkoj logiki//I.V. Anikin/ Sistemnaya inzheneriya i informacionnye tekhnologii, 2023. T. 5, Vyp. 3. № 12. –S. 93–113. DOI 10.54708/2658-5014-SIIT-2023-no3-p93

9. Makarenko S.I. Kriterii i pokazateli ocenki kachestva testirovaniya na proniknovenie//S.I. Makarenko/ Voprosy kiberbezopasnosti, 2021. № 3(43). –S. 43–57. DOI:10.21681/2311-3456-2021-3-43-57

10. Smirnov S.I., Eremeev M.A., Magomedov SH.G., Izergin D.A. Kriterii i pokazateli ocenivaniya kachestva provedeniya rassledovaniya incidenta informacionnoj bezopasnosti pri celevoj kiberatake// S.I. Smirnov, M.A. Eremeev, SH.G. Magomedov, D.A. Izergin /Russian Technological Journal, 2024. T. 12. Vyp. 3. –S. 25–36. DOI: https://doi.org/10.32362/2500-316X-2024-12-3-25-36

11. Makarenko S. I., Smirnov G. E. Metodika obosnovaniya testovyh informacionno-tekhnicheskih vozdejstvij, obespechivayushchih racional'nuyu polnotu audita zashchishchennosti ob"ekta kriticheskoj informacionnoj infrastruktury//S. I. Makarenko, G. E. Smirnov/ Voprosy kiberbezopasnosti. 2021. № 6. (46). –S. 12–25. DOI:10.21681/2311-3456-2021-6-12-25

12. Kotenko I.V. i dr. Tekhnologii bol'shih dannyh dlya korrelyacii sobytij bezopasnosti na osnove ucheta tipov svyazej//I.V. Kotenko, A.V. Fedorchenko, I.B. Saenko, A.G. Kushnerevich A.G./ Voprosy kiberbezopasnosti,YU 2017. №5 (24). –S. 2–16. DOI: 10.21681/2311-3456-2017-5-2-16

13. Organizaciya sistemy setevogo monitoringa i ocenki sostoyaniya informacionnoj bezopasnosti ob"ekta / A.L. Maruhlenko, K.D. Seleznyov, M.O. Tanygin, L.O. Maruhlenko // Izvestiya YUgo-Zapadnogo gosudarstvennogo universiteta, 2019. T. 23. № 1. –S. 118–129. DOI: 10.21869/2223-1560-2019-23-1118-129.

14. Kobzar' A.I. Prikladnaya matematicheskaya statistika. Dlya inzhenerov i nauchnyh rabotnikov. –2-e izd., ispr. –M.: FIZMATLIT, 2012. –816 s.

15. URL: https://CVEform.mitre.org (Data obrashcheniya: 22.12.2025).

16. URL: https://attack.mitre.org/ (Data obrashcheniya: 22.12.2025).

17. URL: https://d3fend.mitre.org/ (Data obrashcheniya: 22.12.2025.

18. Bank dannyh ugroz bezopasnosti informacii// URL: https://bdu.fstec.ru/threat (Data obrashcheniya: 22.12.2025).

19. Metodicheskij dokument «Metodika ocenki ugroz bezopasnosti informacii» (utv. FSTEK 05.02.2021 g.)// URL: https://www.consultant.ru/document/cons_doc_LAW_378330/ (Data obrashcheniya: 22.12.2025).

20. Kocaman Yu., Gönen S., Barişkan M.A., Karacayilmaz G., Yilmaz E.N. A Novel approach to continuous CVE analysis on enterprise operating systems for vulnerability assessment. International journal of information technology (Singapore), 2022. Vol. 14. № 3. P. 1433–1443.

21. Aghaei E, Al-Shaer E, Shadid W, Niu X. Automated cve analysis for threat prioritization and impact prediction. 2023. arXiv preprint arXiv:2309.03040. DOI: https://doi.org/10.48550/arXiv.2309.03040

22. Manjunatha A, Kota K, Babu AS, et al. Cve severity prediction from vulnerability description-a deep learning approach. Procedia Comput Sci, 2024; 235:3105–17. DOI: 10.1016/j.procs.2024.04.294

23. Sonmeza F. O., C. Hankina, Malacariab P. Dynamics: An Automatic Attack Graph Generation Framework Based on System Topology, CAPEC, CWE, and CVE Databases. Preprint submitted to Computers and Security September 26, 2022. 22 p. URL: https://www.researchgate.net/publication/364096733_Attack_Dynamics_An_Automatic_Attack_Graph_Generation_Framework_Based_on_System_Topology_CAPEC_CWE_and_CVE_Database

24. Tulinov I.S., Gubskov YU.A., Gromov YU.YU., SHatskih V.V. Formirovanie bazy dannyh uyazvimostej programmnogo obespecheniya dlya operativnoj nejtralizacii ugroz// I.S. Tulinov, YU.A. Gubskov, YU.YU. Gromov YU.YU., V.V. SHatskih/ Promyshlennye ASU i kontrollery, 2018, № 3. –S. 26–30.

25. Paiva R. Cernambi Virgem Ecológico (CVE), un nuevo caucho de extracción natural con calidad y directo de la selva amazónica a la industria del calzadohttps. DOI://doi.org/10.4067/S0718-07642021000600101

26. Belkin S.A., Belov V.M., Pivkin E.N. Primenenie faktornogo planirovaniya eksperimenta dlya ocenki veroyatnostej ugroz informacionnoj bezopasnosti// Polzunovskij vestnik, 2014. № 2. –S. 232–234.

27. Pletnev P. V., Belov V. M., Zubkov E. V., Kryzhanovskaya O. A. K voprosu ob opredelenii ugroz i riskov informacionnoj bezopasnosti s ispol'zovaniem scenarnogo podhoda i faktornogo planirovaniya eksperimenta// Vestnik SibGUTI. 2016. № 4. –C. 12–18.

28. Markelova E.B., Troeglazova A.V. Ocenka faktorov, okazyvayushchih vliyanie na realizaciyu ugroz informacionnoj bezopasnosti// INTEREKSPO GEO-SIBIR', 2023. T. 6. № 1. –S. 165–170.

29. SHago F.N. Metodika ocenki effektivnosti sistemy menedzhmenta informacionnoj bezopasnosti (SMIB) po vremeni reakcii sistemy na incidenty IB//Nauchno-tekhnicheskij vestnik informacionnyh tekhnologij, mekhaniki i optiki (Scientific and Technical Journal of Information Technologies, Mechanics and Optics), 2014. № 4 (92). –C. 115–123.

30. URL://www.anti-malware.ru/files/27-04-2023/28_s.jpg (Data obrashcheniya: 22.12.2025).

31. URL: https://d3fend.mitre.org/technique/d3f:HostReboot/ (Data obrashcheniya: 21.12.2025).

ПОРШНЕВ Сергей Владимирович, доктор технических наук, профессор, профессор Учебно-научного центра «Информационная безопасность» Федерального государственного автономного образовательного учреждения высшего образования «Уральский федеральный университет им. первого Президента России Б.Н. Ельцина». 620002, г. Екатеринбург, ул. Мира, 32. E-mail: s.v.porshnev@urfu.ru

САФИУЛЛИН Николай Тахирович, кандидат технических наук, доцент Учебно-научного центра «Информационная безопасность» Федерального государственного автономного образовательного учреждения высшего образования «Уральский федеральный университет им. первого Президента России Б.Н. Ельцина». 620002, г. Екатеринбург, ул. Мира, 32. E-mail: n.t.safiullin@urfu.ru

PORSHNEV Sergey Vladimirovich, Doctor of Technical Sciences, Professor, Professor of the Educational and Scientific Center «Information Security» of the Federal State Autonomous

Educational Institution of Higher Education «Ural Federal University named after the first President of Russia B.N. Yeltsin». 620002, Yekaterinburg, st. Mira, 32. E-mail: s.v.porshnev@urfu.ru

SAFIULLIN Nikolai Tahirovich, Candidate of Technical Sciences, Associate Professor of the Educational and Scientific Center «Information Security» of the Federal State Autonomous Educational Institution of Higher Education «Ural Federal University named after the first President of Russia B.N. Yeltsin». 620002, Yekaterinburg, st. Mira, 32. E-mail: n.t.safiullin@urfu.ru

УДК 004.056.53 Вестник УрФО № 2(60) / 2026, с. 47–58

Гладких К.И., Шабалин А.М., Захаров А.А.

DOI: 10.14529/secur260205

МЕТОДОЛОГИЯ ФОРМИРОВАНИЯ КОНТЕКСТА УЯЗВИМОСТЕЙ ДЛЯ ИХ ОБНАРУЖЕНИЯ ЯЗЫКОВЫМИ МОДЕЛЯМИ НА ОСНОВЕ ГРАФА СВОЙСТВ КОДА

В статье рассматривается задача формирования программного контекста для обнаружения уязвимостей исходного кода языковыми моделями. Ограниченность анализа изолированных функций приводит к потере межпроцедурных зависимостей, потоков данных и условий использования, влияющих на проявление уязвимости. Предложена методология, основанная на графе свойств кода, который объединяет структурные элементы программы, граф вызовов и зависимости данных. Методология включает синтаксический разбор Python-проектов, построение узлов функций, классов, переменных и блоков кода, формирование ребер вызовов и потоков данных, загрузку графа в Neo4j, связывание результатов статического анализа с узлами кода, извлечение локального подграфа вокруг потенциально уязвимого участка и ранжирование элементов контекста по графовой удаленности и повторяемости. Полученный контекст адаптируется к ограничению контекстного окна языковой модели. Апробация на 400 экземплярах из набора CVEFixes показала снижение доли ложноотрицательных срабатываний и рост коэффициента Фи для большинства протестированных моделей. Результаты подтверждают, что учет семантического программного контекста повышает устойчивость обнаружения уязвимостей. Работа соответствует пунктам 15 и 17 паспорта научной специальности 2.3.6.

Ключевые слова: *уязвимость, статический анализ, большая языковая модель, граф свойств кода, программный контекст, информационная безопасность.*

A METHODOLOGY FOR CONSTRUCTING VULNERABILITY CONTEXT FOR DETECTION BY LANGUAGE MODELS BASED ON CODE PROPERTY GRAPHS

The article addresses the problem of constructing program context for vulnerability detection in source code using language models. Function-level analysis often loses interprocedural dependencies, data flows, and usage conditions that determine whether a vulnerability can manifest. The proposed methodology is based on a Code Property Graph that combines program structure, call relations, and data-flow dependencies. The methodology includes parsing Python projects, constructing nodes for functions, classes, variables, and code blocks, creating call-graph and data-flow edges, loading the graph into Neo4j, linking static-analysis findings with code nodes, extracting a local subgraph around a potentially vulnerable fragment, and ranking context elements by graph distance and repetition frequency. The resulting context is adapted to the context-window limit of a language model. The evaluation on 400 CVEFixes-based instances showed a reduction in false-negative predictions and an increase in the Phi coefficient for most tested models. The results indicate that semantically grounded program context improves the robustness of language-model-based vulnerability detection.

Keywords: *vulnerability, static analysis, large language model, code property graph, program context, information security.*

Введение

Большие языковые модели активно применяются для анализа исходного кода [1,2] однако их результативность в задачах поиска уязвимостей зависит не только от архитектуры модели и промпта, но и от состава программного контекста. В распространенной постановке задача сводится к бинарной классификации отдельной функции или небольшого фрагмента. Такая постановка удобна для бенчмарков, но часто не отражает природу уязвимости: источник данных, цепочка вызовов, состояние объекта и условие выполнения могут находиться вне анализируемой функции [3–6].

Отсутствие релевантного контекста приводит к двум типам ошибок. Во-первых, модель может не обнаружить уязвимость, если необходимый источник данных или обработчик расположен в другой части проекта. Во-вторых, расширение контекста без приоритизации увеличивает шум, расходует контекстное окно и может ухудшать классификацию. Поэтому ключевой задачей является не простое добавление фрагментов кода, а выбор компактного множества элементов, непосредственно связанных с потенциально уязвимым участком [7–10].

В литературе выделяются три группы решений. Первая группа использует retrieval-augmented generation, при котором в запрос к модели добавляются сведения об известных уязвимостях, исправлениях или связанных фрагментах репозитория [8,9]. Вторая группа объединяет LLM со статическим анализом, программными срезами и графовыми представлениями кода [7,11–13]. Третья группа применяет агентные архитектуры, где мо-

дель последовательно вызывает инструменты анализа, уточняет гипотезы и формирует итоговый ответ [14,15]. Эти подходы полезны, но имеют ограничения: retrieval может не учитывать фактические зависимости внутри проекта, графовые конвейеры часто требуют ресурсоемкой обработки, а агентные системы зависят от стоимости, времени выполнения и недетерминированности модели.

Дополнительная методологическая проблема связана с качеством наборов данных. Многие датасеты формируются по коммитам исправления уязвимостей и содержат только измененные строки или функции. Это упрощает разметку, но снижает полноту программного контекста и может создавать статистические артефакты, по которым модель учится отличать исправленный код от уязвимого без анализа семантики программы [16–21].

Проблема, цель и задачи исследования

Проблема исследования состоит в том, что существующие способы подачи кода в языковую модель либо анализируют слишком локальные фрагменты, либо добавляют избыточный контекст без явной связи с проявлением уязвимости. В результате модель получает неполную или зашумленную информацию, а качество обнаружения остается нестабильным.

Цель исследования – разработать и экспериментально оценить методологию формирования компактного программного контекста потенциально уязвимых участков кода на основе графа свойств кода, обеспечивающую повышение качества обнаружения уязвимостей языковыми моделями по сравнению с анализом изолированных фрагментов.

Для достижения цели были поставлены следующие задачи:

• проанализировать существующие подходы к формированию контекста для LLM-обнаружения уязвимостей и определить их ограничения;

• разработать способ построения графа свойств кода для Python-проектов на основе синтаксического разбора и межпроцедурных связей;

• предложить алгоритм извлечения и ранжирования контекстного подграфа по релевантности;

• реализовать сборку программного контекста с учетом ограничения контекстного окна модели;

• провести апробацию на данных CVEFixes и оценить статистическую значимость различий между базовым и контекстным подходами.

Методология формирования программного контекста

Объектом анализа является исходный код программного проекта на языке Python. Предметом анализа являются связи между потенциально уязвимым участком и другими элементами программы: вызовами функций, передачей аргументов, определениями переменных, принадлежностью элементов классам и блокам кода. Область исследования ограничена статическим анализом исходного кода и бинарной классификацией экземпляра как уязвимого или неуязвимого. Предполагается, что потенциально уязвимые участки могут быть локализованы либо по строкам исправления в CVE-записях, либо по находкам статических анализаторов.

Методология строится как последовательный конвейер. Сначала проект разбирается синтаксическим анализатором. Затем строится граф свойств кода. После этого результаты статического анализа связываются с узлами графа. Далее вокруг корневых узлов извлекается локальный подграф, его узлы ранжируются, а соответствующие фрагменты исходного кода добавляются в промпт до достижения лимита контекстного окна. Общая схема представлена на рис. 1.

Построение графа свойств кода

Граф свойств кода (Code Property Graph, CPG) представляет программу как граф, в котором синтаксические элементы, потоки управления и зависимости данных описываются в единой структуре [21]. В данной работе CPG адаптирован для Python-кода. Синтаксическая структура извлекается с помощью tree-sitter-python; затем из дерева формируются узлы и ребра, отражающие семантически значимые отношения между частями программы.

В граф включаются четыре типа узлов. FunctionNode описывает функцию или метод, содержит имя, файл и диапазон строк, а также связывается с параметрами. ClassNode описывает класс и его тело. VariableNode фиксирует параметры, присваивания и переменные текущей области видимости. CodeBlockNode представляет непрерывный блок императивных операторов уровня модуля и используется как вызывающий контекст для скриптового кода.

Ребра графа делятся на две основные груп-

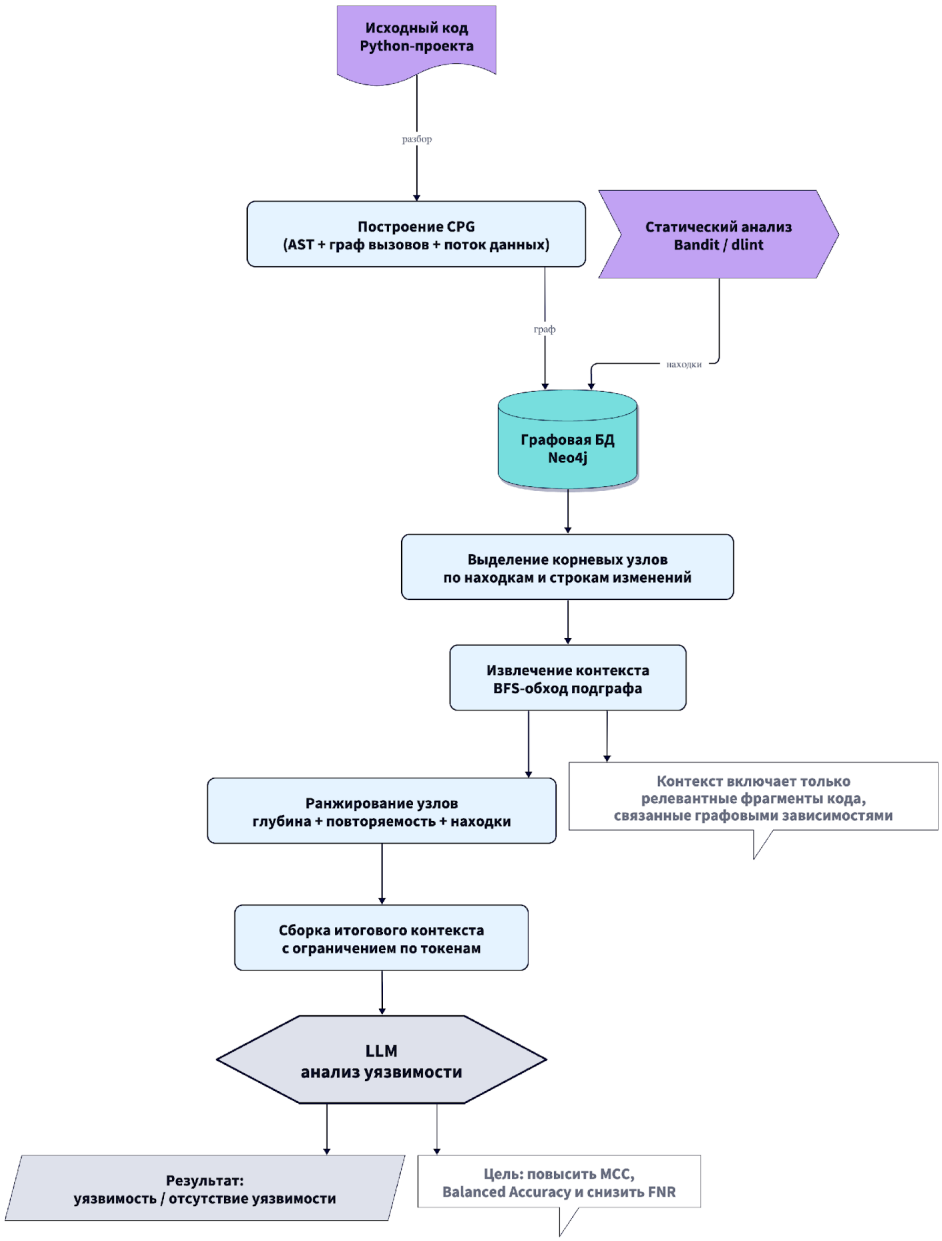

Рис. 1. Схема формирования программного контекста уязвимости

пы. Ребра CallGraphCalls и CallGraphCalledBy фиксируют вызовы функций и методов. Если вызываемый объект является идентификатором, он разрешается по цепочке областей видимости; если это атрибут, разрешается имя метода. Ребра DataFlowDefinedBy и DataFlowFlowsTo описывают определение значения и передачу данных в аргументы вызовов. Такая модель позволяет извлекать не только локальный фрагмент, но и окружающие его зависимости.

Построение выполняется в несколько проходов. На уровне модуля сначала выделяются блоки императивного кода, затем предварительно связываются классы и функции, после чего выполняется полный рекурсивный обход синтаксического дерева. Для разрешения идентификаторов используется стек областей видимости, а для атрибуции вызовов – стек вызывающего контекста. При обработке каталога сначала формируется индекс экспортируемых имен по файлам, затем выполняется повторная сборка с учетом межмодульных импортов.

Практическая реализация разделена на три компонента. Класс CPGFileBuilder обрабатывает один исходный файл: считывает байты, декодирует текст в UTF-8, строит конкрет-

ное синтаксическое дерево с помощью tree_sitter.Parser и нормализует путь относительно корня проекта. CPGDirectoryBuilder распространяет обработку на дерево каталогов, формирует индекс экспортируемых имен и повторно строит граф с учетом межмодульных импортов. NodeProcessor выполняет рекурсивный обход дерева, поддерживая стек областей видимости и стек вызывающего контекста.

Текущая реализация имеет ограничения. Управляющие конструкции не выделяются в самостоятельный граф потока управления, а учитываются через кодовые узлы, в которые уже входит соответствующий синтаксический фрагмент. Разрешение вызовов методов выполняется по имени метода и не моделирует все случаи динамической диспетчеризации Python. Эти ограничения учитывались при интерпретации результатов.

Хранение графа и извлечение подграфа

Для хранения CPG используется графовая база Neo4j. Элементы исходного кода загружаются пакетно: каждая запись сопоставляется с узлом Code и дополнительной меткой, уточняющей его роль, например, Function, Class, Variable или Call. Для связей используется параметризованная загрузка Cypher-запросами. Применение MERGE обеспечивает идемпотентное создание узлов и снижает риск дублирования при повторной обработке проекта.

Контекст извлекается как локальный подграф вокруг множества корневых узлов. Корневым считается узел, связанный с потенциально уязвимым фрагментом. Запрос выполняет обход графа на глубину N по выбранным типам ребер и возвращает идентификатор корня, идентификатор найденного узла, путь к файлу, диапазон строк, имя, тип узла и длину пути. Длина пути используется как оценка графовой удаленности.

Пример контекстного подграфа, построенного для функции main при глубине 2, показан на рис. 2.

Определение потенциально уязвимых участков

Потенциально уязвимый участок кода определяется как фрагмент, который взаимодействует с внешними или частично доверенными источниками данных, ресурсами либо подсистемами: базами данных, файловой системой, сетевыми интерфейсами, пользовательским вводом или системными командами. Ошибка в таком фрагменте может нарушить конфиденциальность, целостность или доступность данных.

В экспериментальной реализации такие

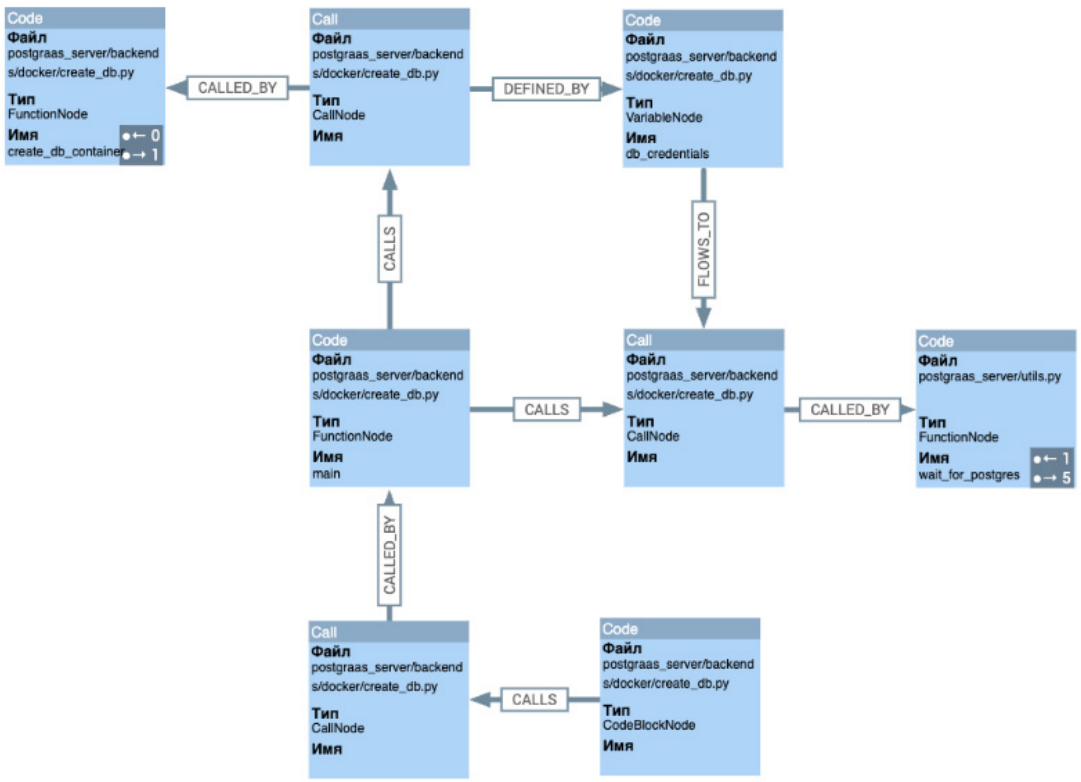

Рис. 2. Пример результата построения контекстного подграфа для функции main с глубиной 2

участки выявляются двумя способами. В режиме апробации на CVEFixes корневые узлы определяются по файлам и строкам, затронутым исправлением уязвимости. В режиме анализа проекта корневые узлы могут формироваться по находкам статических анализаторов Bandit и dlint. Каждая находка сохраняется в графе и связывается с соответствующим узлом ребром StaticAnalysisReports.

Ранжирование и сбор контекста

Так как контекстное окно языковой модели ограничено, весь подграф не передается в модель. Узлы ранжируются по трем признакам: графовой глубине относительно корня, числу повторений в множестве подграфов и числу связанных находок статического анализа. В базовой реализации приоритет получает узел с меньшей глубиной, большей повторяемостью и большим числом находок.

После ранжирования исходный код узлов последовательно добавляется в результирующий контекст. Добавление прекращается, когда расчетное число токенов достигает заданного лимита. Оценка числа токенов выполняется библиотекой tiktoken для целевого типа языковой модели. Таким образом, методология не увеличивает контекст бесконтрольно, а использует граф как механизм отбора и упорядочивания программной информации.

Апробация методологии

Апробация выполнена на базе CVEFixes – набора данных, содержащего реальные CVE-записи, связанные репозитории и коммиты исправления уязвимостей [22]. Для каждой записи извлекалось состояние репозитория до исправления и после исправления. Затем для обоих состояний строился CPG, данные загружались в Neo4j, а по файлам и строкам исправления определялись корневые узлы.

Для получения контекста применялся обход графа в ширину на глубину 6. Из найденных подграфов извлекались уникальные узлы, подсчитывалась их повторяемость, затем узлы сортировались по возрастанию глубины, убыванию числа повторений и числу находок. Соответствующие фрагменты кода объединялись в экземпляр контекста до достижения лимита контекстного окна.

В итоговую выборку вошло 400 экземпляров исходного кода. Часть записей была исключена, так как результирующий код превышал допустимый размер контекстного окна или не мог быть корректно сопоставлен с узлами графа. Для сравнения использовались два варианта подачи данных: исходный фрагмент CVEFixes и контекст, сформированный предложенным методом ContextAssembler.

В экспериментах применялись модели Gemma3 4B, Llama 3.2 3B, Qwen3-4B thinking, Qwen3-8B thinking, Qwen3.5-4B и Qwen3.5-4B thinking. Температура была равна 0,3. Размер контекстного окна составлял 24 000 токенов для моделей с режимом рассуждения и 20 000 токенов для остальных моделей. Во всех запусках использовался одинаковый seed. Промпт к моделям представлен на рис. 3.

Оценка выполнялась по общей доле правильных классификаций (accuracy), полноте (recall), специфичности (specificity), доле ложноотрицательных срабатываний (FNR), сбалансированной точности и коэффициенту

```
You are a methodical security analyst who follows a systematic
approach to vulnerability detection.
Analysis Steps:
1. Read through the code to understand its purpose
2. Identify all input sources and data flows
3. Check for proper input validation and bounds checking
4. Look for memory management issues
5. Examine error handling and edge cases
6. Determine if vulnerabilities exist\n\nFinal Answer: After
your analysis, respond with \"VULNERABLE\" or \"SAFE\"
```

Рис. 3. Промпт к языковым моделям для анализа исходного кода на предмет уязвимостей

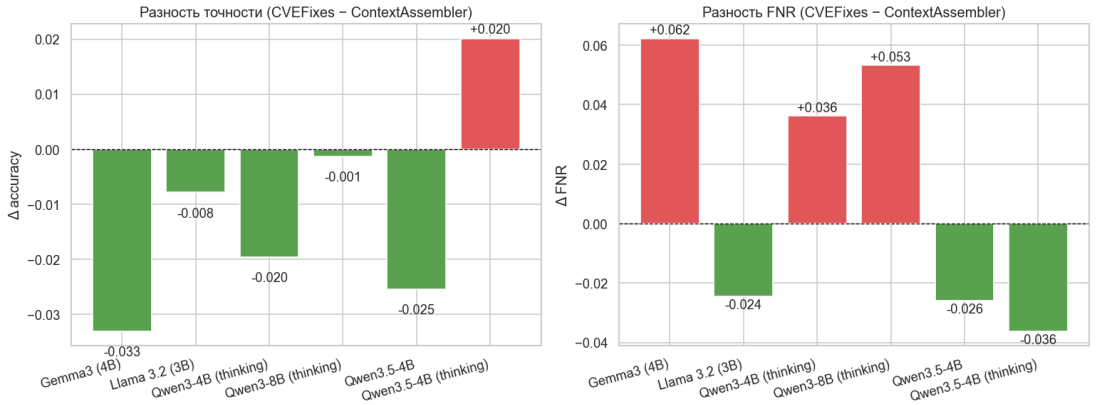

Рис. 4. Сравнение общей доли правильных классификаций и ложноотрицательных срабатываний моделей на разных наборах данных

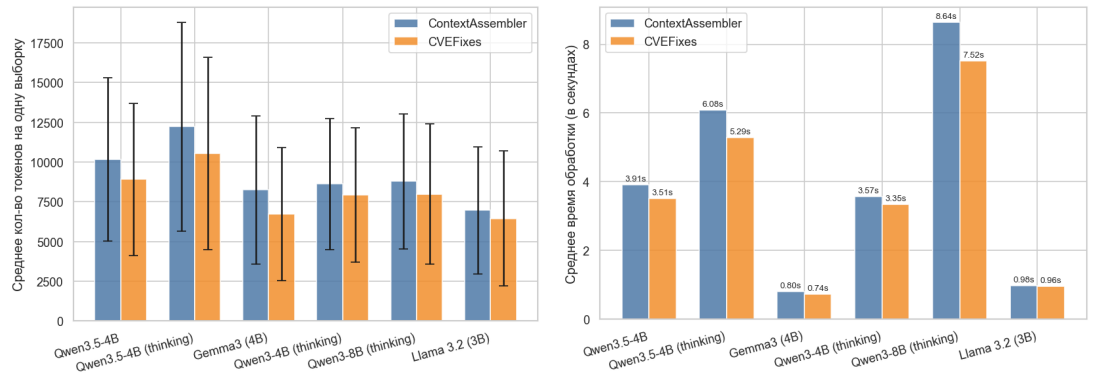

Рис. 5. Средние показатели потребления токенов и времени обработки одного экземпляра кода

корреляции Фи. Для проверки статистической значимости различий между двумя способами формирования контекста применялся тест Макнемара с поправкой на непрерывность. Предсказания агрегировались на уровне CVE-записей методом голосования большинством.

Результаты

На рис. 4 показано сравнение accuracy и доли ложноотрицательных срабатываний. Для большинства моделей ContextAssembler снижает FNR по сравнению с исходным представлением CVEFixes. Это особенно важно для задач безопасности, поскольку ложноотрицательный ответ означает пропуск уязвимости. При этом рост accuracy не является универсальным: для отдельных моделей наблюдаются небольшие отклонения, что указывает на зависимость эффекта от архитектуры и режима генерации.

На рис. 5 приведены средние показатели потребления токенов и времени обработки одного экземпляра. Увеличение числа токенов является ожидаемым, так как методология добавляет межпроцедурный контекст. Однако время обработки для большинства моделей изменилось умеренно; наиболее заметные отклонения наблюдались у семейства Qwen.

На рис. 6 показано сравнение коэффициента Фи и сбалансированной точности. Значения коэффициента Фи для ContextAssembler положительны во всех экспериментах, тогда как для базового CVEFixes в отдельных случаях они близки к нулю или отрицательны. Это означает, что контекстное представление уменьшает случайность классификации и повышает согласованность предсказаний с истинной разметкой.

Статистическая проверка показала, что для четырех из шести моделей различия между подходами значимы при $p < 0{,}01$. Для Gemma3 4B, Llama 3.2 3B, Qwen3-8B thinking и Qwen3.5-4B thinking ContextAssembler давал больше случаев, в которых контекстный подход был корректен, а базовый подход ошибался. Для Qwen3-4B thinking и Qwen3.5-4B статистическая значимость не достигнута. Результаты приведены в табл. 1.

Сводные метрики приведены в табл. 2. По

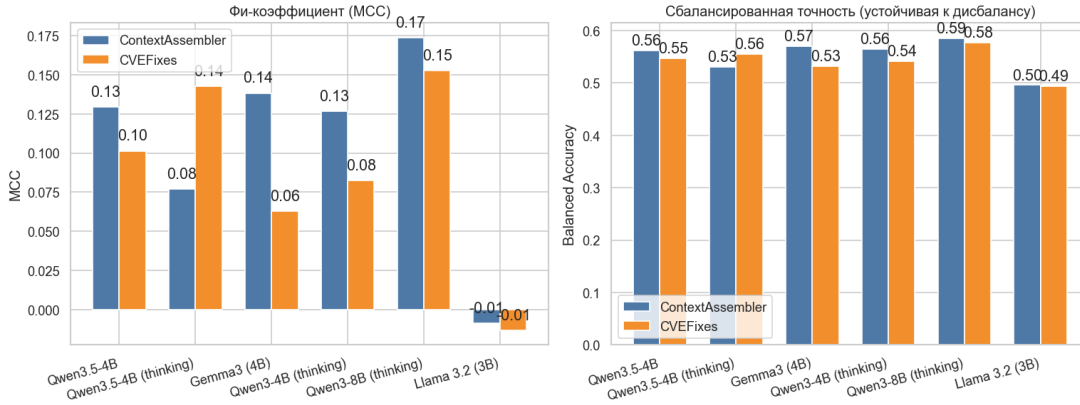

Рис. 6. Сравнение моделей по коэффициенту корреляции Фи и сбалансированной точности

Таблица 1

Оценка статистической значимости тестом Макнемара

Модель	Совпадение по CVE	Оба корректны	CA ✓ CVE ✗	CA ✗ CVE ✓	Оба некорректны	χ^2	p-value	Значимо
Gemma3 (4B)	120	20	37	16	47	7,547	0,0060	да
Llama 3.2 (3B)	131	12	41	18	60	8,203	0,0042	да
Qwen3-4B (thinking)	128	29	31	19	49	2,420	0,1198	нет
Qwen3-8B (thinking)	128	27	39	14	48	10,868	0,0010	да
Qwen3.5-4B	124	22	32	21	49	1,887	0,1696	нет
Qwen3.5-4B (thinking)	124	20	37	11	56	13,021	0,0003	да

сравнению с CVEFixes предложенный подход увеличивает коэффициент Фи для Gemma3 4B, Llama 3.2 3B, Qwen3-4B thinking, Qwen3-8B thinking и Qwen3.5-4B. Для Qwen3.5-4B thinking коэффициент Фи ниже, однако FNR также уменьшается, что показывает смещение модели в сторону более чувствительного обнаружения уязвимостей.

Дальнейшие направления для исследования

Полученные результаты показывают, что графовое формирование контекста особенно полезно для снижения ложноотрицательных срабатываний. Это соответствует предположению исследования: уязвимость часто определяется не только локальным кодом, но и связанными с ним источниками данных, вызовами и определениями. Одновременно результаты не позволяют утверждать, что методология всегда повышает accuracy. На отдельных моделях эффект выражается иначе, что требует дополнительного анализа промптов, архитектур и стратегии ранжирования.

Основное ограничение работы связано с упрощенным ранжированием контекста. В текущей версии используются глубина, повторяемость и число находок, однако эти признаки не учитывают тип уязвимости, направление потока данных, доверенность источника и наличие санитизации. Второе ограничение состоит в неполной модели потока управления для Python. Третье ограничение связано с размером выборки: 400 экземпляров достаточно для первичной апробации, но не заменяет крупномасштабную оценку на разных языках программирования и типах проектов.

Дальнейшее развитие методологии пред-

Таблица 2

Полные результаты бенчмарка

Набор данных	Модель	Количество	Accuracy	Recall	Specificity	FNR	Фи
Context-Assembler	Gemma3 (4B)	394	0,551	0,667	0,472	0,333	0,138
	Llama 3.2 (3B)	400	0,455	0,700	0,292	0,300	–0,009
	Qwen3-4B (thinking)	400	0,565	0,562	0,567	0,438	0,127
	Qwen3-8B (thinking)	400	0,555	0,738	0,433	0,262	0,174
	Qwen3.5-4B	400	0,528	0,738	0,388	0,262	0,130
	Qwen3.5-4B (thinking)	400	0,465	0,856	0,204	0,144	0,077
CVEFixes	Gemma3 (4B)	394	0,518	0,604	0,459	0,396	0,063
	Llama 3.2 (3B)	400	0,447	0,724	0,264	0,276	–0,013
	Qwen3-4B (thinking)	400	0,545	0,526	0,558	0,474	0,082
	Qwen3-8B (thinking)	400	0,554	0,684	0,469	0,316	0,153
	Qwen3.5-4B	400	0,502	0,763	0,331	0,237	0,101
	Qwen3.5-4B (thinking)	400	0,485	0,892	0,218	0,108	0,143

полагает включение taint-анализа, учет типов CWE и CVE-метаданных, более строгую модель межпроцедурного потока данных, а также адаптацию промптов под конкретные классы уязвимостей. Отдельным направлением является оценка устойчивости метода на репозиторном уровне, где необходимо учитывать зависимости между файлами, внешними библиотеками и конфигурацией проекта.

Заключение

В статье предложена методология формирования программного контекста потенциально уязвимых участков кода на основе графа свойств кода. Методология объединяет синтаксическое представление программы, граф вызовов, потоки данных и результаты статического анализа, после чего извлекает локальный подграф вокруг корневого узла и преобразует его в компактный текстовый контекст для языковой модели.

Экспериментальная апробация на данных CVEFixes показала, что контекст, сформированный с помощью графа свойств кода, снижает долю ложноотрицательных срабатываний и повышает коэффициент Фи для большинства протестированных моделей. Статистический анализ тестом Макнемара подтвердил значимый сдвиг предсказаний для четырех из шести моделей. Следовательно, учет межпроцедурного и семантического контекста является существенным условием повышения надежности LLM-систем, применяемых для анализа безопасности исходного кода.

Литература

1. Мельников, А. В. Автоматизация оценки вредоносности скриптов на основе больших языковых моделей при расследовании компьютерных инцидентов / А. В. Мельников, И. Н. Ярышкин // Вестник УрФО. Безопасность в информационной сфере. – 2025. – № 4(58). – С. 36-43. – DOI 10.14529/secur250404. – EDN RWTFCR.

2. Нейронные сети для обеспечения безопасности веб-приложений / В. А. Частикова, А. А. Тесленко, М. К. Алиев, И. С. Игнатенко // Вестник УрФО. Безопасность в информационной сфере. – 2025. – № 2(56). – С. 65-73. – DOI 10.14529/secur250207. – EDN UYYZGA.

3. Risse, N. Top Score on the Wrong Exam: On Benchmarking in Machine Learning for Vulnerability Detection / N. Risse, J. Liu, M. Bohme. – Текст: непосредственный // ArXiv. – 2024. – № abs/2408.12986. – URL: https://arxiv.org/abs/2408.12986 (дата обращения: 01.04.2026).

4. Ding, Y. Vulnerability Detection with Code Language Models: How Far Are We? / Y. Ding, Y. Fu, O. Ibrahim [et al.]. – Текст: непосредственный // Proceedings of the 2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE). – 2025. – С. 1729–1741. – DOI: 10.1109/ICSE55347.2025.00038.

5. Zhang, Y. Context and Multi-Features-Based Vulnerability Detection: A Vulnerability Detection Frame Based on Context Slicing and Multi-Features / Y. Zhang, Y. Hu, X. Chen. – Текст: непосредственный // Sensors. – 2024. – Т. 24, № 5. – Ст. 1351. – DOI: 10.3390/s24051351.

6. Li, Z. IRIS: LLM-Assisted Static Analysis for Detecting Security Vulnerabilities / Z. Li, S. Dutta, M. Naik. – Текст: непосредственный // ArXiv. – 2024. – № abs/2405.17238. – URL: https://arxiv.org/abs/2405.17238 (дата обращения: 02.04.2026).

7. Du, X. Vul-RAG: Enhancing LLM-Based Vulnerability Detection via Knowledge-Level RAG / X. Du, G. Zheng, K. Wang [et al.]. – Текст: непосредственный // ArXiv. – 2024. – № abs/2406.11147. – URL: https://arxiv.org/abs/2406.11147 (дата обращения: 07.04.2026).

8. Wen, X.-C. VulEval: Towards Repository-Level Evaluation of Software Vulnerability Detection / X.-C. Wen, X. Wang, Y. Chen [et al.]. – Текст: непосредственный // ArXiv. – 2024. – № abs/2404.15596. – URL: https://arxiv.org/abs/2404.15596 (дата обращения: 07.04.2026).

9. Li, Y. Beyond Function-Level Analysis: Context-Aware Reasoning for Inter-Procedural Vulnerability Detection / Y. Li, T. Zhang, J. Shi [et al.]. – Текст: непосредственный // ArXiv. – 2026. – № abs/2602.06751. – URL: https://arxiv.org/abs/2602.06751 (дата обращения: 10.04.2026).

10. Lekssays, A. LLMxCPG: Context-Aware Vulnerability Detection Through Code Property Graph-Guided Large Language Models / A. Lekssays, H. Mouhcine, K. Tran [et al.]. – Текст: непосредственный // ArXiv. – 2025. – № abs/2507.16585. – URL: https://arxiv.org/abs/2507.16585 (дата обращения: 10.04.2026).

11. Wang, C. Boosting Static Resource Leak Detection via LLM-Based Resource-Oriented Intention Inference / C. Wang, J. Liu, X. Peng, Y. Liu, Y. Lou. – Текст: непосредственный // ArXiv. – 2023. – № abs/2311.04448. – URL: https://arxiv.org/abs/2311.04448 (дата обращения: 10.04.2026).

12. Chapman, P. J. Interleaving Static Analysis and LLM Prompting / P. J. Chapman, C. Rubio-Gonzalez, A. V. Thakur. – Текст: непосредственный // Proceedings of the 13th ACM SIGPLAN International Workshop on the State of the Art in Program Analysis (SOAP). – New York: ACM, 2024. – С. 9–17. – DOI: 10.1145/3652588.3663317.

13. Yildiz, A. Benchmarking LLMs and LLM-Based Agents in Practical Vulnerability Detection for Code Repositories / A. Yildiz, S. G. Teo, Y. Lou [et al.]. – Текст: непосредственный // ArXiv. – 2025. – № abs/2503.03586. – URL: https://arxiv.org/abs/2503.03586 (дата обращения: 14.04.2026).

14. Guo, J. RepoAudit: An Autonomous LLM-Agent for Repository-Level Code Auditing / J. Guo, C. Wang, X. Xu, Z. Su, X. Zhang. – Текст: непосредственный // Proceedings of the 42nd International Conference on Machine Learning. – PMLR, 2025. – Т. 267. – С. 21083–21100.

15. Li, Y. CleanVul: Automatic Function-Level Vulnerability Detection in Code Commits Using LLM Heuristics / Y. Li, T. Zhang, R. Widyasari [et al.]. – Текст: непосредственный // ArXiv. – 2024. – № abs/2411.17274. – URL: https://arxiv.org/abs/2411.17274 (дата обращения: 17.04.2026).

16. Ahmed, M. B. U. SecVulEval: Benchmarking LLMs for Real-World C/C++ Vulnerability Detection / M. B. U. Ahmed, N. S. Harzevili, J. Shin, H. V. Pham, S. Wang. – Текст: непосредственный // ArXiv. – 2025. – № abs/2505.19828. – URL: https://arxiv.org/abs/2505.19828 (дата обращения: 17.04.2026).

17. Li, Yue. Everything You Wanted to Know About LLM-Based Vulnerability Detection But Were Afraid to Ask / Yue Li, Xiao Li, H. Wu [et al.]. – Текст: непосредственный // ArXiv. – 2025. – № abs/2504.13474. – URL: https://arxiv.org/abs/2504.13474 (дата обращения: 16.04.2026).

18. Zhang, X. GitPatchDB: A Large-Scale GitHub Commit Databank for Vulnerability Patch Analysis / X. Zhang, P. He, X. Jin, Y. Kao, S. Chen. – Текст: непосредственный // OpenReview. – 2025. – URL: https://openreview.net/forum?id=tX24AtTHyg (дата обращения: 16.04.2026).

19. Wang, X. ReposVul: A Repository-Level High-Quality Vulnerability Dataset / X. Wang, R. Hu, C. Gao [et al.]. – Текст : непосредственный // ArXiv. – 2024. – № abs/2401.13169. – URL: https://arxiv.org/abs/2401.13169 (дата обращения: 17.04.2026).

20. Guo, Y. A Comprehensive Analysis on Software Vulnerability Detection Datasets: Trends, Challenges, and Road Ahead / Y. Guo, S. Bettaieb, F. Casino. – Текст : непосредственный // International Journal of Information Security. – 2024. – Т. 23, № 5. – С. 3311–3327. – DOI: 10.1007/s10207-024-00888-y.

21. Yamaguchi, F. Modeling and Discovering Vulnerabilities with Code Property Graphs / F. Yamaguchi, N. Golde, D. Arp, K. Rieck. – Текст: непосредственный // 2014 IEEE Symposium on Security and Privacy. – San Jose: IEEE, 2014. – С. 590–604. – DOI: 10.1109/SP.2014.44.

22. Bhandari, G. P. CVEfixes: Automated Collection of Vulnerabilities and Their Fixes from Open-Source Software / G. P. Bhandari, A. Naseer, L. Moonen. – Текст: непосредственный // Proceedings of the 17th International Conference on Predictive Models and Data Analytics in Software Engineering (PROMISE 2021). – New York: ACM, 2021. – Ст. 8. – 10 с. – DOI: 10.1145/3475960.3475985.

References

1. Mel'nikov, A. V. Avtomatizaciya ocenki vredonosnosti skriptov na osnove bol'shih yazykovyh modelej pri rassledovanii komp'yuternyh inci-dentov / A. V. Mel'nikov, I. N. YAryshkin // Vestnik UrFO. Bezopasnost' v informacionnoj sfere. – 2025. – № 4(58). – S. 36-43. – DOI 10.14529/secur250404. – EDN RWTFCR.

2. Nejronnye seti dlya obespecheniya bezopasnosti veb-prilozhenij / V. A. CHastikova, A. A. Teslenko, M. K. Aliev, I. S. Ignatenko // Vestnik UrFO. Bezopasnost' v informacionnoj sfere. – 2025. – № 2(56). – S. 65–73. – DOI 10.14529/secur250207. – EDN UYYZGA.

3. Risse N., Liu J., Bohme M. Top Score on the Wrong Exam: On Benchmarking in Machine Learning for Vulnerability Detection. arXiv:2408.12986, 2024. Available at: https://arxiv.org/abs/2408.12986

4. Ding Y., Fu Y., Ibrahim O., Sitawarin C., Chen X., Alomair B., Wagner D., Ray B., Chen Y. Vulnerability Detection with Code Language Models: How Far Are We? Proceedings of the 2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE). 2025, pp. 1729–1741. DOI: 10.1109/ICSE55347.2025.00038

5. Zhang Y., Hu Y., Chen X. Context and Multi-Features-Based Vulnerability Detection: A Vulnerability Detection Frame Based on Context Slicing and Multi-Features. Sensors, 2024, vol. 24, no. 5, article 1351. DOI: 10.3390/s24051351

6. Li Z., Dutta S., Naik M. IRIS: LLM-Assisted Static Analysis for Detecting Security Vulnerabilities. arXiv:2405.17238, 2024. Available at: https://arxiv.org/abs/2405.17238

7. Du X., Zheng G., Wang K., Feng J., Deng W., Liu M., Chen B., Peng X., Ma T., Lou Y. Vul-RAG: Enhancing LLM-Based Vulnerability Detection via Knowledge-Level RAG. arXiv:2406.11147, 2024. Available at: https://arxiv.org/abs/2406.11147

8. Wen X.-C., Wang X., Chen Y., Hu R., Lo D., Gao C. VulEval: Towards Repository-Level Evaluation of Software Vulnerability Detection. arXiv:2404.15596, 2024. Available at: https://arxiv.org/abs/2404.15596

9. Li Y., Zhang T., Shi J., Yang C., He J., Zhou X., Jiang J., Huang H., Leow W.B., Yin Y., Ouh E.L., Shar L.K., Lo D. Beyond Function-Level Analysis: Context-Aware Reasoning for Inter-Procedural Vulnerability Detection. arXiv:2602.06751, 2026. Available at: https://arxiv.org/abs/2602.06751

10. Lekssays A., Mouhcine H., Tran K., Yu T., Khalil I. LLMxCPG: Context-Aware Vulnerability Detection Through Code Property Graph-Guided Large Language Models. arXiv:2507.16585, 2025. Available at: https://arxiv.org/abs/2507.16585

11. Wang C., Liu J., Peng X., Liu Y., Lou Y. Boosting Static Resource Leak Detection via LLM-Based Resource-Oriented Intention Inference. arXiv:2311.04448, 2023. Available at: https://arxiv.org/abs/2311.04448

12. Chapman P.J., Rubio-Gonzalez C., Thakur A.V. Interleaving Static Analysis and LLM Prompting. Proceedings of the 13th ACM SIGPLAN International Workshop on the State of the Art in Program Analysis (SOAP). New York, ACM, 2024, pp. 9–17. DOI: 10.1145/3652588.3663317

13. Yildiz A., Teo S.G., Lou Y., Feng Y., Wang C., Divakaran D.M. Benchmarking LLMs and LLM-Based Agents in Practical Vulnerability Detection for Code Repositories. arXiv:2503.03586, 2025. Available at: https://arxiv.org/abs/2503.03586

14. Guo J., Wang C., Xu X., Su Z., Zhang X. RepoAudit: An Autonomous LLM-Agent for Repository-Level Code Auditing. Proceedings of the 42nd International Conference on Machine Learning. PMLR, 2025, vol. 267, pp. 21083–21100

15. Li Y., Zhang T., Widyasari R., Tun Y.N., Nguyen H.H., Bui T., Irsan I.C., Cheng Y., Lan X., Ang H.W., Liauw F., Weyssow M., Kang H.J., Ouh E.L., Shar L.K., Lo D. CleanVul: Automatic Function-Level Vulnerability Detection in Code Commits Using LLM Heuristics. arXiv:2411.17274, 2024. Available at: https://arxiv.org/abs/2411.17274

16. Ahmed M.B.U., Harzevili N.S., Shin J., Pham H.V., Wang S. SecVulEval: Benchmarking LLMs for Real-World C/C++ Vulnerability Detection. arXiv:2505.19828, 2025. Available at: https://arxiv.org/abs/2505.19828

17. Yue Li, Xiao Li, Wu H., Xu M., Zhang Y., Cheng X., Xu F., Zhong S. Everything You Wanted to Know About LLM-Based Vulnerability Detection But Were Afraid to Ask. arXiv:2504.13474, 2025. Available at: https://arxiv.org/abs/2504.13474

18. Zhang X., He P., Jin X., Kao Y., Chen S. GitPatchDB: A Large-Scale GitHub Commit Databank for Vulnerability Patch Analysis. OpenReview, 2025. Available at: https://openreview.net/forum?id=tX24AtTHyg

19. Wang X., Hu R., Gao C., Wen X.-C., Chen Y., Liao Q. ReposVul: A Repository-Level High-Quality Vulnerability Dataset. arXiv:2401.13169, 2024. Available at: https://arxiv.org/abs/2401.13169

20. Guo Y., Bettaieb S., Casino F. A Comprehensive Analysis on Software Vulnerability Detection Datasets: Trends, Challenges, and Road Ahead. International Journal of Information Security, 2024, vol. 23, no. 5, pp. 3311–3327. DOI: 10.1007/s10207-024-00888-y

21. Yamaguchi F., Golde N., Arp D., Rieck K. Modeling and Discovering Vulnerabilities with Code Property Graphs. 2014 IEEE Symposium on Security and Privacy. San Jose, IEEE, 2014, pp. 590–604. DOI: 10.1109/SP.2014.44

22. Bhandari G.P., Naseer A., Moonen L. CVEfixes: Automated Collection of Vulnerabilities and Their Fixes from Open-Source Software. Proceedings of the 17th International Conference on Predictive Models and Data Analytics in Software Engineering (PROMISE 2021). New York, ACM, 2021, article 8, 10 p. DOI: 10.1145/3475960.3475985

ГЛАДКИХ Кирилл Игоревич, аспирант, Школа компьютерных наук, Тюменский государственный университет. 625003. г. Тюмень, ул. Володарского, д. 6. E-mail: boombarah@gmail.com.

ШАБАЛИН Андрей Михайлович, кандидат педагогических наук, доцент, доцент академического департамента, Школа компьютерных наук, Тюменский государственный университет. 625003. г. Тюмень, ул. Володарского, д. 6. E-mail: sham.omsk@gmail.com.

ЗАХАРОВ Александр Анатольевич, доктор технических наук, профессор, Школа компьютерных наук, Тюменский государственный университет. 625003. г. Тюмень, ул. Володарского, д. 6. E-mail: a.a.zakharov@utmn.ru.

GLADKIKH Kirill Igorevich, post-graduate student, School of Computer Science, Tyumen State University. 625003, Tyumen, Volodarsky St., 6. E-mail: boombarah@gmail.com.

SHABALIN Andrey Mihailovich, Candidate of Pedagogical Sciences, Associate Professor, Associate Professor of the Academic Department, School of Computer Science, Tyumen State University. 625003, Tyumen, Volodarsky St., 6. E-mail: sham.omsk@gmail.com.

ZAKHAROV Alexander Anatolievich, Doctor of Technical Sciences, Professor, School of Computer Science, Tyumen State University. 625003, Tyumen, Volodarsky St., 6. E-mail: a.a.zakharov@utmn.ru.

УДК 004.056 Вестник УрФО № 2(60) / 2026, с. 59–67

Лукоянов А.А.

DOI: 10.14529/secur260206

МОДЕЛЬ СБОРА ВЕРИФИЦИРОВАННЫХ ДАННЫХ ДЛЯ ИНТЕГРАЛЬНОЙ ОЦЕНКИ ЗАЩИЩЕННОСТИ РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ

В статье рассматривается задача сбора и проверки исходных данных, используемых для интегральной оценки защищенности распределенной информационной системы. Достоверность итогового показателя зависит не только от выбранной математической модели, но и от качества данных, поступающих из средств мониторинга информационной безопасности, систем инвентаризации, средств анализа уязвимостей, подсистем управления доступом, журналов приложений, средств резервного копирования и систем регистрации инцидентов. В работе обоснована необходимость отдельной проверки исходных данных до их использования в автоматизированном расчете. Проверка включает оценку полноты, актуальности, согласованности, доверия к источнику и целостности данных. Предложена модель сбора верифицированных данных и введен показатель качества исходной информационной базы, который используется совместно с интегральным показателем защищенности.

Ключевые слова: *информационная безопасность, распределенная информационная система, интегральная оценка, мониторинг, верификация данных, качество данных.*

Lukoyanov A.A.

MODEL FOR COLLECTING VERIFIED DATA FOR INTEGRAL SECURITY ASSESSMENT OF A DISTRIBUTED INFORMATION SYSTEM

The paper considers the problem of collecting and verifying source data used for integral security assessment of a distributed information system. It is shown that the reliability of the fi-

nal security indicator depends not only on the selected mathematical model, but also on the quality of data received from information security monitoring tools, asset inventory systems, vulnerability scanners, access management subsystems, application logs, backup systems and incident registration systems. The need to introduce a separate verification layer is substantiated. This layer checks completeness, relevance, consistency, source trustworthiness and data integrity before the data are admitted to calculation. A verified data collection model is proposed, a data verification algorithm is developed, and a source data quality indicator is introduced for joint use with the integral security indicator.

Keywords: *information security, distributed information system, integral assessment, monitoring, data verification, data quality.*

Введение. Интегральная оценка защищенности информационной системы используется для обобщенного представления состояния защиты, сопоставления объектов, контроля динамики изменений и поддержки управленческих решений. Для распределенных информационных систем такая оценка важна, поскольку прикладные сервисы, базы данных, сетевые компоненты, средства защиты, пользовательские рабочие места и внешние интеграции могут находиться в различных сегментах инфраструктуры.

Расчет интегрального показателя рассматривается как преимущественно математическая задача: выбирается набор частных показателей, каждому показателю назначается вес, после чего выполняется агрегирование. Достоверность результата зависит не только от выбранной расчетной формулы, но и от качества исходных данных: их полноты, актуальности, взаимной согласованности и доверия к используемым источникам.

В работе О.В. Казарина, С.Е. Кондакова и И.И. Троицкого обоснован переход от качественного описания состояния защищенности информационных систем к его количественной оценке [1]. С.Е. Кондаков отдельно рассматривал вопросы количественной оценки защищенности информации от несанкционированного доступа [2]. Д.А. Рыленков и Д.С. Карпов предложили интегральный метод оценки уровня защищенности серверной инфраструктуры организации [3], а Е.В. Бурькова исследовала особенности оценки защищенности информационных систем персональных данных [4]. В работах И.И. Лившица и А.С. Бакшеева данная проблематика рассматривается применительно к контролю уровня защищенности информации на объектах критической информационной инфраструктуры [5, 6]. В документах NIST SP 800-55, NIST SP 800-137, NIST Cybersecurity Framework 2.0 и ISO/IEC 27004 отмечается, что метрики безопасности должны быть связаны с источниками данных, процессами мониторинга, оценкой состояния защиты и принятием управленческих решений [7–9, 13, 14]. ГОСТ Р 59547-2021 также определяет мониторинг информационной безопасности как процесс получения и анализа данных, характеризующих состояние защищенности [10]. Современные исследования показывают, что при анализе уязвимостей и оценке защищенности необходимо учитывать не только значения отдельных показателей, но и качество данных, на которых основан расчет [15, 17]. Отдельные работы посвящены количественной и интегральной оценке защищенности компонентов информационной инфраструктуры [3, 16].

Существующие подходы к построению метрик информационной безопасности в основном ориентированы на выбор показателей, их измерение и последующее агрегирование, тогда как предварительная проверка качества исходных данных перед расчетом интегрального показателя защищенности рассматривается недостаточно подробно. Для распределенных информационных систем эта проблема имеет особое значение, поскольку данные поступают из разных источников: SIEM, CMDB, сканеров уязвимостей, IAM, EDR/XDR, систем резервного копирования и журналов приложений. Эти источники различаются форматами, периодичностью обновления, задержками передачи данных и степенью доверия. В результате без отдельной проверки в расчет могут попадать неполные, устаревшие или несогласованные данные, а также записи, которые невозможно надежно соотнести с конкретным сервером, сервисом, учетной записью или сетевым сегментом.

Целью статьи является разработка модели и алгоритма сбора верифицированных данных для интегральной оценки защищенности распределенной информационной системы.

Научная новизна. Научная новизна исследования состоит в том, что в интегральную оценку защищенности распределенной информационной системы включена предварительная проверка качества исходных данных. В отличие от подходов, где основное внимание уделяется выбору показателей и формуле их объединения, в данной работе отдельно рассматривается достоверность данных, поступающих из разных источников. Предложенная модель проверяет данные по пяти критериям: полнота, актуальность, согласованность, доверие к источнику и целостность. Результат оценки предлагается представлять в виде двух показателей: уровня защищенности и качества данных, на которых основан расчет.

Постановка задачи. Распределенную информационную систему представим как множество активов:

$$A = \{a_1, a_2, \ldots, a_n\}, \quad (1)$$

где A_i – отдельный актив или компонент системы: сервер, рабочая станция, сетевое устройство, база данных, прикладной сервис, средство защиты, API-компонент, учетная запись или сегмент сети.

Для каждого актива могут оцениваться несколько направлений защищенности:

$$R = \{r_1, r_2, \ldots, r_m\}, \quad (2)$$

где r_j – направление оценки: управление доступом, управление уязвимостями, защищенность конфигураций, мониторинг событий информационной безопасности, реагирование на инциденты, резервное копирование и восстановление, защита внешних взаимодействий.

Первичная запись, поступающая от средств мониторинга или управления информационной безопасностью, может быть представлена в виде:

$$D = <a, s, t, p, v, r>, \quad (3)$$

где a – актив, к которому относится запись; s – источник данных; t – время получения данных; p – измеряемый параметр; v – значение параметра; r – направление оценки защищенности. Такая запись не должна автоматически включаться в расчет, поскольку она может быть устаревшей, неполной, противоречивой или неподтвержденной.

Задача состоит в преобразовании первичных записей в верифицированные данные, пригодные для расчета частных и интегрального показателей защищенности.

Состав данных для оценки защищенности. Для интегральной оценки защищенности распределенной информационной системы необходимо учитывать события информационной безопасности, данные о составе активов, уязвимостях, конфигурациях, доступах, инцидентах, резервном копировании и внешних взаимодействиях. События информационной безопасности отражают зарегистрированные факты функционирования системы, а данные об активах, настройках и уязвимостях позволяют учитывать потенциальные слабые места ее защиты.

Для распределенной системы принципиальное значение имеет связь данных с акти-

Таблица 1

Основные группы данных для оценки защищенности распределенной ИС

Группа данных	Возможные источники	Применение к оценке
Состав активов	CMDB, инвентаризация активов, сетевое сканирование	Определение границ системы и полноты охвата
Инциденты	IRP/SOAR, журналы регистрации инцидентов, отчеты реагирования	Оценка выявления, классификации и устранения инцидентов
Уязвимости	Сканеры уязвимостей, отчеты анализа защищенности	Оценка технической уязвимости компонентов
Конфигурации	Средства контроля конфигураций, эталонные профили	Оценка соответствия безопасным настройкам
Доступы	IAM, AD/LDAP, журналы привилегированных действий	Оценка управления доступом
События информационной безопасности	SIEM, EDR/XDR, IDS/IPS, WAF	Оценка мониторинга, обнаружения и реагирования
Восстановление	Backup-системы, отчеты тестового восстановления	Оценка устойчивости восстановления

вами. Значение показателя, не привязанное к конкретному серверу, сервису, учетной записи или сетевому сегменту, трудно использовать в расчете. Поэтому модель должна обеспечивать нормализацию идентификаторов и сопоставление данных из разных источников с единым перечнем активов.

Модель сбора верифицированных данных. Предлагаемая модель технологического контура сбора и обработки данных включает шесть функциональных уровней: уровень источников данных, уровень сбора и доставки, уровень нормализации, уровень привязки к активам и направлениям оценки, уровень верификации данных и уровень расчета показателей защищенности.

Общая структура модели представлена на рис. 1.

Рис. 1. Модель сбора верифицированных данных для оценки защищенности распределенной информационной системы

Первичные данные о состоянии системы поступают из средств мониторинга, инвентаризации, анализа уязвимостей, управления доступом и иных подсистем. На уровне сбора и доставки обеспечивается их получение с использованием агентов, коннекторов, API, Syslog, файловых выгрузок или регламентированных отчетов. На этапе нормализации данные преобразуются к единой структуре представления, что позволяет сопоставлять записи из разных источников. В предлагаемой модели исходные данные не передаются непосредственно в расчетный модуль. Их предварительная проверка выполняется в самостоятельном контуре верификации, который обеспечивает исключение из расчетной выборки записей, не отвечающих установленным требованиям к полноте, актуальности, согласованности, доверию к источнику и целостности.

Критерии верификации данных. Для оценки качества исходных данных предлагается использовать пять критериев: полнота, актуальность, согласованность, доверие к источнику и целостность. Выбор этих критериев согласуется с общими подходами к оценке качества данных, в которых подчеркивается значение полноты, точности, своевременности и соответствия данных задачам последующего анализа [11, 12].

Коэффициент полноты показывает, достаточно ли данных для расчета:

$$C = \frac{N_{пол}}{N_{тр}}, \qquad (4)$$

где $N_{пол}$ – количество полученных обязательных значений; $N_{тр}$ – количество значений, необходимых для расчета.

Коэффициент актуальности характеризует соответствие данных установленному интервалу их обновления:

$$F = \max\left(0, 1 - \frac{\Delta t}{T_{доп}}\right), 0 \le F \le 1, \qquad (5)$$

где Δt – время, прошедшее с момента получения или последнего обновления данных; $T_{доп}$ – максимально допустимый интервал времени для соответствующего типа данных.

Все частные коэффициенты качества исходных данных принимают значения на отрезке [0; 1]. Если данные устарели настолько, что значение Δt превышает $T_{доп}$, коэффициент актуальности принимается равным нулю.

Коэффициент согласованности опреде-

Таблица 2

Критерии верификации исходных данных

Критерий	Обозначение	Содержание	Пример проверки
Полнота	C	Наличие необходимых данных по активам и направлениям	Все ли серверы охвачены сканированием
Актуальность	F	Данные получены или обновлены в пределах установленного интервала времени	Не устарел ли отчет сканера
Согласованность	G	Отсутствие существенных противоречий между источниками	Совпадает ли статус актива в CMDB и SIEM
Доверие к источнику	S	Уровень доверия к источнику и способу получения данных	Источник входит в утвержденный перечень доверенных источников данных
Целостность	I	Контроль целостности данных	Используется ли защищенный канал или контрольная сумма

ляется как доля контрольных сопоставлений, по результатам которых не выявлены противоречия между данными из разных источников:

$$G = 1 - \frac{N_{пр}}{N_{пров}}, \quad (6)$$

где $N_{пр}$ – количество выявленных противоречий; $N_{пров}$ – количество выполненных проверок согласованности. При этом $N_{пров} > 0$.

Коэффициент доверия к источнику S задается на основе утвержденного перечня источников данных. Коэффициент целостности I принимает значение 1 при подтвержденной целостности, 0,5 при частичном подтверждении и 0 при отсутствии подтверждения.

Итоговый показатель качества исходных данных рассчитывается по формуле:

$$Q_D = C \cdot F \cdot G \cdot S \cdot I; \quad (7)$$

Мультипликативная форма агрегирования позволяет ограничить взаимное замещение критериев: высокое значение одного критерия качества данных не должно полностью устранять влияние существенных отклонений по другим критериям, включая полноту, доверие к источнику и целостность данных.

Параметры модели задаются на основании регламентов эксплуатации информационной системы, состава контролируемых активов, утвержденных источников данных и требований к периодичности мониторинга. Количество обязательных значений для расчета коэффициента полноты определяется перечнем данных, необходимых для соответствующего направления оценки. Максимально допустимый интервал актуальности устанавливается с учетом типа данных и частоты их обновления: для событий информационной безопасности – по времени поступления в SIEM, для данных об уязвимостях – по периодичности сканирования, для данных об активах – по регламенту актуализации CMDB, для резервного копирования – по периодичности формирования отчетов и проверок восстановления.

Значение коэффициента доверия к источнику определяется по утвержденной шкале: 1 – основной доверенный источник, данные которого формируются автоматически и передаются по защищенному каналу; 0,75 – дополнительный доверенный источник; 0,5 – косвенный или периодически актуализируемый источник; 0 – источник, не включенный в утвержденный перечень или не обеспечивающий подтверждение происхождения данных. Коэффициент целостности определяется по результатам проверки отсутствия искажений при передаче, хранении и обработке данных. Минимально допустимый порог качества данных определяется владельцем информационной системы или подразделением информационной безопасности с учетом допустимого риска принятия решений на основе неполных, устаревших или противоречивых данных. Весовые коэффициенты направлений оценки устанавливаются с учетом критичности соответствующих направлений защищенности и нормируются таким образом, чтобы их сумма была равна единице.

Алгоритм сбора и проверки данных. Предлагаемый алгоритм описывает последовательность преобразования первичных

данных мониторинга в набор верифицированных данных, используемых при расчете интегрального показателя защищенности. Общая логика алгоритма представлена на рис. 2. На первом этапе формируется множество активов распределенной информационной системы и определяются направления оценки. Для каждого направления назначаются источники данных: SIEM, EDR/XDR, сканеры уязвимостей, системы управления учетными записями и правами доступа, CMDB, системы резервного копирования и журналы прикладных сервисов. Полученные данные приводятся к единому формату: актив, источник, время, параметр, значение и направление оценки. Записи сопоставляются с активами из сформированного перечня. Данные, не имеющие связи с конкретным активом, направляются на дополнительную проверку и не включаются в расчет. Ключевым этапом является расчет обобщенного показателя качества данных Q_D и его сравнение с минимально допустимым порогом Q_{min}. При выполнении условия $Q_D \geq Q_{min}$ данные включаются в верифицированный набор и используются для расчета частных показателей и интегрального показателя защищенности $K_{зи}$. Если условие не выполняется, данные уточняются, собираются повторно либо исключаются из расчетной выборки.

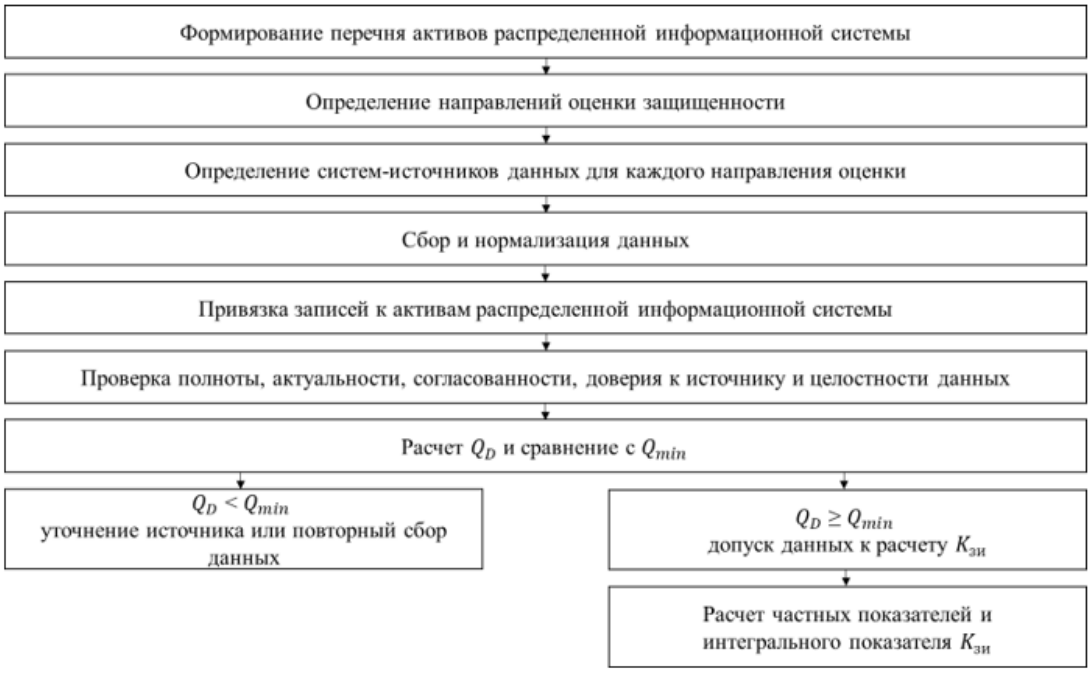

Рис. 2. Алгоритм сбора и верификации данных для расчета интегрального показателя защищенности

Использование верифицированных данных при интегральной оценке. Пусть частные показатели защищенности по направлениям оценки имеют вид K_1, K_2, ..., K_m. Для их агрегирования может использоваться взвешенное геометрическое среднее, позволяющее учитывать существенное снижение защищенности по отдельному направлению даже при высоких значениях остальных частных показателей:

$$K_{зи} = \prod_{j=1}^{m} K_j^{w_j}, \quad (8)$$

где $K_{зи}$ – интегральный показатель защищенности; K_j – частный показатель по j-му направлению; w_j — вес j-го направления, при этом сумма весов равна 1.

Каждому направлению оценки может быть сопоставлен собственный показатель качества данных Q_{Dj}. Используются те же веса w_j, что и для направлений защищенности. Обобщенный показатель качества исходных данных определяется аналогично:

$$Q_{Dобщ} = \prod_{j=1}^{m} Q_{Dj}^{w_j}; \quad (9)$$

Результат оценки предлагается представлять в виде пары:

$$<K_{зи}; Q_{Dобщ}>; \quad (10)$$

Первый элемент характеризует расчетный уровень защищенности, второй – качество исходных данных, на которых основан расчет. В отдельных случаях может использоваться скорректированная оценка защищенности, учитывающая качество исходных данных:

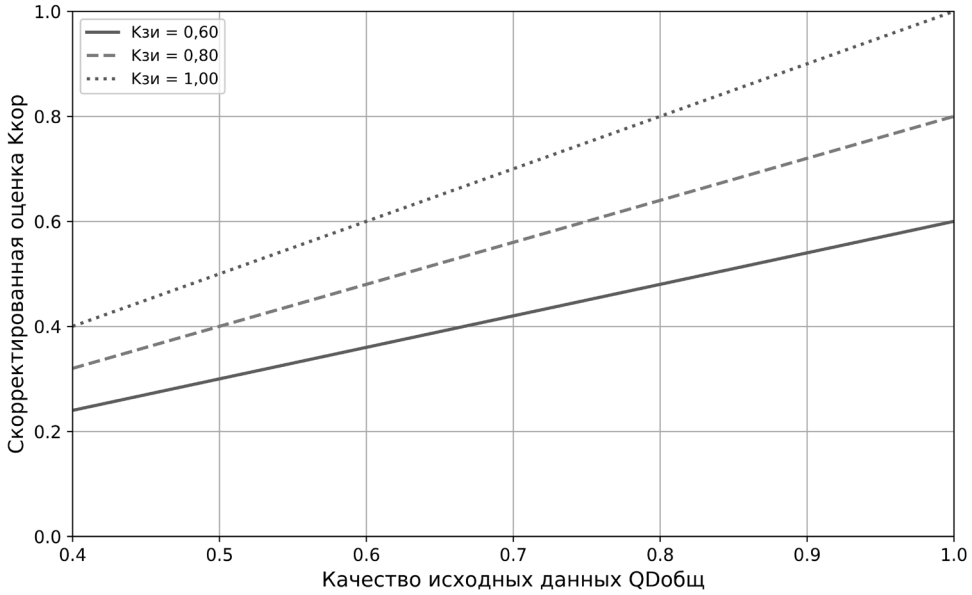

Рис. 3. Зависимость скорректированной оценки защищенности от качества исходных данных

$$K_{кор} = K_{зи} \cdot Q_{Добщ}; \quad (11)$$

Показатель $K_{кор}$ не заменяет $K_{зи}$, а служит для управленческой интерпретации результата, когда необходимо учитывать риск принятия решения на основе неполной или недостаточно проверенной информации.

На рис. 3 показана зависимость скорректированной оценки защищенности от качества исходных данных при нескольких фиксированных значениях расчетного показателя защищенности.

Пример расчета. Рассмотрим пример, в котором оцениваются четыре направления защищенности: управление уязвимостями, управление доступом, мониторинг событий ИБ, резервное копирование и восстановление.

Таблица 3

Пример оценки защищенности и качества исходных данных

Направление оценки	K_j	Q_{Dj}	Вес w_j
Управление уязвимостями	0,72	0,80	0,30
Управление доступами	0,85	0,90	0,25
Мониторинг событий информационной безопасности	0,78	0,70	0,25
Резервное копирование и восстановление	0,88	0,75	0,20

По формуле (8) интегральный показатель защищенности составляет $K_{зи} \approx 0{,}80$. По формуле (9) обобщенный показатель качества исходных данных составляет $Q_{Добщ} \approx 0{,}79$. Следовательно, результат оценки может быть представлен как <0,80; 0,79>.

Если используется скорректированная оценка защищенности, то по формуле (11) получаем $K_{кор} = 0{,}80 \cdot 0{,}79 = 0{,}63$. Следовательно, без учета качества исходных данных система могла бы быть охарактеризована как имеющая достаточно высокий уровень защищенности, однако с учетом качества данных скорректированная оценка защищенности оказывается ниже. Это указывает на необходимость повышения полноты, актуальности или согласованности исходной информационной базы.

Заключение. Предложена модель сбора верифицированных данных для интегральной оценки защищенности распределенной информационной системы. В отличие от подходов, ориентированных преимущественно на выбор частных показателей и способ их агрегирования, модель включает самостоятельный контур проверки качества исходных данных до выполнения расчета.

Разработан алгоритм сбора и верифика-

ции данных, предусматривающий формирование перечня активов, определение направлений оценки, назначение источников, нормализацию, привязку записей к активам, проверку полноты, актуальности, согласованности, доверия к источнику и целостности данных. Предложено представлять результат оценки в виде пары $<K_{зи}; Q_{Добщ}>$, где $K_{зи}$ характеризует расчетный уровень защищенности, а $Q_{Добщ}$ отражает качество исходных данных. Такой подход повышает интерпретируемость результата и снижает риск принятия управленческих решений на основе неполных, устаревших или противоречивых данных.

Практическая значимость работы заключается в возможности применения модели при построении подсистем мониторинга информационной безопасности, автоматизированной оценке защищенности, подготовке отчетности для владельца информационной системы и контроле качества данных, используемых для принятия решений в области защиты информации.

Литература

1. Казарин О.В., Кондаков С.Е., Троицкий И.И. Подходы к количественной оценке защищенности ресурсов автоматизированных систем // Вопросы кибербезопасности. 2015. № 2(10). С. 31–35.

2. Кондаков С.Е. К вопросу о количественной оценке защищенности информации от несанкционированного доступа в информационных системах // Вопросы кибербезопасности. 2015. № 4(12). С. 28–35.

3. Рыленков Д.А., Карпов Д.С. Разработка интегрального метода оценки уровня защищенности серверной инфраструктуры организации // Инженерный вестник Дона. 2025. № 1.

4. Бурькова Е.В. Задача оценки защищенности информационных систем персональных данных // Вестник Чувашского университета. 2016. № 1. С. 112–118.

5. Лившиц И.И., Бакшеев А.С. Исследование методик контроля уровня защищенности информации на объектах критической информационной инфраструктуры // Вопросы кибербезопасности. 2022. № 6(52). С. 40–52. DOI: 10.21681/2311-3456-2022-6-40-52.

6. Бакшеев А.С., Лившиц И.И. Разработка методики контроля уровня защищенности информации объектов критической информационной инфраструктуры // Вопросы кибербезопасности. 2023. № 2(54). С. 85–98. DOI: 10.21681/2311-3456-2023-2-85-98.

7. Schroeder K., Trinh H., Pillitteri V.Y. Measurement Guide for Information Security: Volume 1 – Identifying and Selecting Measures. NIST Special Publication 800-55 Vol. 1. National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.SP.800-55v1.

8. Dempsey K., Chawla N.S., Johnson A., Johnston R., Jones A.C., Orebaugh A., Scholl M., Stine K. Information Security Continuous Monitoring for Federal Information Systems and Organizations. NIST Special Publication 800-137. National Institute of Standards and Technology, 2011. DOI: 10.6028/NIST.SP.800-137.

9. ISO/IEC 27004:2016. Information technology — Security techniques — Information security management — Monitoring, measurement, analysis and evaluation.

10. ГОСТ Р 59547-2021. Защита информации. Мониторинг информационной безопасности. Общие положения.

11. Wang R.Y., Strong D.M. Beyond Accuracy: What Data Quality Means to Data Consumers // Journal of Management Information Systems. 1996. Vol. 12. No. 4. P. 5–33. DOI: 10.1080/07421222.1996.11518099.

12. Pipino L.L., Lee Y.W., Wang R.Y. Data Quality Assessment // Communications of the ACM. 2002. Vol. 45. No. 4. P. 211–218. DOI: 10.1145/505248.506010.

13. Schroeder K., Trinh H., Pillitteri V.Y. Measurement Guide for Information Security: Volume 2 – Developing a Measurement Program. NIST Special Publication 800-55 Vol. 2. National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.SP.800-55v2.

14. National Institute of Standards and Technology. The NIST Cybersecurity Framework (CSF) 2.0. NIST Cybersecurity White Paper 29. National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.CSWP.29.

15. Croft R., Babar M.A., Kholoosi M.M. Data Quality for Software Vulnerability Datasets // Proceedings of the 45th International Conference on Software Engineering. 2023. P. 121–133. DOI: 10.1109/ICSE48619.2023.00022.

16. Янгиров А.И., Янгиров И.М., Рогозин Е.А., Ахлюстин С.Б. Методический подход к количественной оценке защищенности открытых операционных систем АС ОВД // Вестник Дагестанского государственного технического университета. Технические науки. 2024. Т. 51, № 3. С. 72–85.

17. Кучкарова Н.В. Оценка актуальных угроз и уязвимостей объектов критической информаци-

онной инфраструктуры с использованием технологий интеллектуального анализа текстов // СИИТ. 2024. Т. 6, № 2(17). С. 50–65. DOI: 10.54708/2658-5014-SIIT-2024-no2-p50. EDN NLDWBE.

References

1. Kazarin O.V., Kondakov S.E., Troitskii I.I. Podkhody k kolichestvennoy otsenke zashchishchennosti resursov avtomatizirovannykh sistem // Voprosy kiberbezopasnosti. 2015. № 2(10). S. 31–35.

2. Kondakov S.E. K voprosu o kolichestvennoy otsenke zashchishchennosti informatsii ot nesanktsionirovannogo dostupa v informatsionnykh sistemakh // Voprosy kiberbezopasnosti. 2015. № 4(12). S. 28–35.

3. Rylenkov D.A., Karpov D.S. Razrabotka integral'nogo metoda otsenki urovnya zashchishchennosti servernoy infrastruktury organizatsii // Inzhenernyy vestnik Dona. 2025. № 1.

4. Burkova E.V. Zadacha otsenki zashchishchennosti informatsionnykh sistem personal'nykh dannykh // Vestnik Chuvashskogo universiteta. 2016. № 1. S. 112–118.

5. Livshits I.I., Baksheev A.S. Issledovanie metodik kontrolya urovnya zashchishchennosti informatsii na obektakh kriticheskoy informatsionnoy infrastruktury // Voprosy kiberbezopasnosti. 2022. № 6(52). S. 40–52. DOI: 10.21681/2311-3456-2022-6-40-52.

6. Baksheev A.S., Livshits I.I. Razrabotka metodiki kontrolya urovnya zashchishchennosti informatsii obektov kriticheskoy informatsionnoy infrastruktury // Voprosy kiberbezopasnosti. 2023. № 2(54). S. 85–98. DOI: 10.21681/2311-3456-2023-2-85-98.

7. Schroeder K., Trinh H., Pillitteri V.Y. Measurement Guide for Information Security: Volume 1 – Identifying and Selecting Measures. NIST Special Publication 800-55 Vol. 1. National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.SP.800-55v1.

8. Dempsey K., Chawla N.S., Johnson A., Johnston R., Jones A.C., Orebaugh A., Scholl M., Stine K. Information Security Continuous Monitoring for Federal Information Systems and Organizations. NIST Special Publication 800-137. National Institute of Standards and Technology, 2011. DOI: 10.6028/NIST.SP.800-137.

9. ISO/IEC 27004:2016. Information Technology — Security Techniques — Information Security Management — Monitoring, Measurement, Analysis and Evaluation.

10. GOST R 59547-2021. Zashchita informatsii. Monitoring informatsionnoy bezopasnosti. Obshchie polozheniya.

11. Wang R.Y., Strong D.M. Beyond Accuracy: What Data Quality Means to Data Consumers // Journal of Management Information Systems. 1996. Vol. 12. No. 4. P. 5–33. DOI: 10.1080/07421222.1996.11518099.

12. Pipino L.L., Lee Y.W., Wang R.Y. Data Quality Assessment // Communications of the ACM. 2002. Vol. 45. No. 4. P. 211–218. DOI: 10.1145/505248.506010.

13. Schroeder K., Trinh H., Pillitteri V.Y. Measurement Guide for Information Security: Volume 2 – Developing a Measurement Program. NIST Special Publication 800-55 Vol. 2. National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.SP.800-55v2.

14. National Institute of Standards and Technology. The NIST Cybersecurity Framework (CSF) 2.0. NIST Cybersecurity White Paper 29. National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.CSWP.29.

15. Croft R., Babar M.A., Kholoosi M.M. Data Quality for Software Vulnerability Datasets // Proceedings of the 45th International Conference on Software Engineering. 2023. P. 121–133. DOI: 10.1109/ICSE48619.2023.00022.

16. Yangirov A.I., Yangirov I.M., Rogozin E.A., Akhlyustin S.B. Metodicheskiy podkhod k kolichestvennoy otsenke zashchishchennosti otkrytykh operatsionnykh sistem AS OVD // Vestnik Dagestanskogo gosudarstvennogo tekhnicheskogo universiteta. Tekhnicheskie nauki. 2024. T. 51, № 3. S. 72–85.

17. Kuchkarova N.V. Otsenka aktual'nykh ugroz i uyazvimostey obektov kriticheskoy informatsionnoy infrastruktury s ispol'zovaniem tekhnologiy intellektual'nogo analiza tekstov // SIIT. 2024. T. 6, № 2(17). S. 50–65. DOI: 10.54708/2658-5014-SIIT-2024-no2-p50. EDN NLDWBE.

ЛУКОЯНОВ Алексей Александрович, специалист по информационной безопасности. RUSSPASS — Московская цифровая туристская платформа. Автономная некоммерческая организация «Проектный офис по развитию туризма и гостеприимства Москвы». 125009, г. Москва, ул. Тверская, д. 5А. E-mail: 9261781233@mail.ru

LUKOYANOV Aleksey Aleksandrovich, Information Security Specialist, RUSSPASS — Moscow Digital Tourism Platform. Autonomous Non-profit Organization "Project Office for the Development of Tourism and Hospitality of Moscow". 125009, Moscow, 5A Tverskaya St. E-mail: 9261781233@mail.ru.

УДК 004.056.5 Вестник УрФО № 2(60) / 2026, с. 68–75

Никонов В.В., Трофимов И.А.

DOI: 10.14529/secur260207

МНОГОУРОВНЕВАЯ ЗАЩИТА КОРПОРАТИВНЫХ СЕТЕЙ: МОДЕЛИ МОНИТОРИНГА И ОРГАНИЗАЦИОННО-ТЕХНИЧЕСКИЕ МЕРЫ ПРОТИВОДЕЙСТВИЯ УГРОЗАМ

Целью исследования является разработка методологических основ и организационно-технического комплекса для непрерывного контроля, раннего выявления и нейтрализации современных киберугроз. На базе системного анализа векторов компрометации (DNS-спуфинг, уязвимости нулевого дня, эксплуатация IoT-сегмента) предложены модели мониторинга сетевого трафика, алгоритмы внедрения криптографической верификации DNS-записей и архитектура многоуровневой защиты. В результате сформирован структурированный регламент реагирования на инциденты и разработана матрица практических мероприятий с распределением зон ответственности, периодичностью контроля и механизмами аудита. Обоснована эффективность интеграции поведенческого анализа, автоматизированных систем обнаружения аномалий и криптографических протоколов для минимизации поверхности атаки. Практическая значимость работы заключается в возможности адаптации предложенного комплекса для внедрения в политики информационной безопасности организаций, что способствует повышению операционной устойчивости, снижению рисков утечки данных и соответствию актуальным нормативным требованиям.

Ключевые слова: *кибербезопасность, методы мониторинга, DNS spoofing, атаки нулевого дня, IoT-устройства, защита сети, системы обнаружения и предотвращения вторжений, многоуровневая система безопасности, информационная безопасность, защита данных, угрозы в компьютерных сетях.*

MULTI-LEVEL PROTECTION OF CORPORATE NETWORKS: MONITORING MODELS AND ORGANIZATIONAL AND TECHNICAL MEASURES TO COUNTER THREATS

The aim of this study is to develop a methodological framework and organizational and technical framework for continuous monitoring, early detection, and mitigation of modern cyberthreats. Based on a systemic analysis of compromise vectors (DNS spoofing, zero-day vulnerabilities, IoT segment exploitation), network traffic monitoring models, algorithms for implementing cryptographic verification of DNS records, and a multi-layered defense architecture are proposed. As a result, a structured incident response procedure has been developed and a matrix of practical measures has been developed, assigning responsibilities, monitoring frequency, and audit mechanisms. The effectiveness of integrating behavioral analysis, automated anomaly detection systems, and cryptographic protocols to minimize the attack surface is substantiated. The practical significance of this work lies in the adaptability of the proposed framework for implementation into organizations' information security policies, which contributes to increased operational resilience, reduced risks of data leakage, and compliance with current regulatory requirements.

Keywords: *cybersecurity, monitoring methods, DNS spoofing, zero-day attacks, IoT devices, network protection, intrusion detection and prevention systems, multi-layered security, information security, data protection, threats in computer networks.*

Введение

В настоящее время в цифровой среде компьютерные сети сталкиваются с множеством опасных киберугроз, которые постоянно совершенствуются и для того, чтобы минимизировать риск атаки на системы необходимо методы защиты и повышать уровень кибербезопасности. Одним из таких видов атак являются DNS spoofing или отравление DNS-кэша.

Данный вид атаки реализуется, когда DNS кэш заполняется фальшивыми данными и в результате чего пользователь попадает на вредоносный сайт. Когда происходит отравление DNS-кэша сервер кэша DNS сохраняет поддельный адрес, который предоставил злоумышленник пользователю, а затем передает его пользователям, которые запрашивают настоящий веб-сайт. Главная опасность DNS спуфинга заключается в том, что его очень трудно обнаружить и он может оказать отрицательное влияние на востребованные веб-сайты с большим количеством зарегистрированных пользователей. Это создает огромный риск для таких сфер, как банковская, электронная коммерция и т. д.

Например, предположим, что злоумышленнику удается изменить DNS-записи и IP-адреса для организации. После этих действий они направляют свой запрос на соседний сервер с поддельным IP-адресом, который контролирует злоумышленник. Пользователь, который захочет получить доступ к настоящему сайту организации, будет перенаправлен на поддельный адрес, который содержит вредоносное программное обеспе-

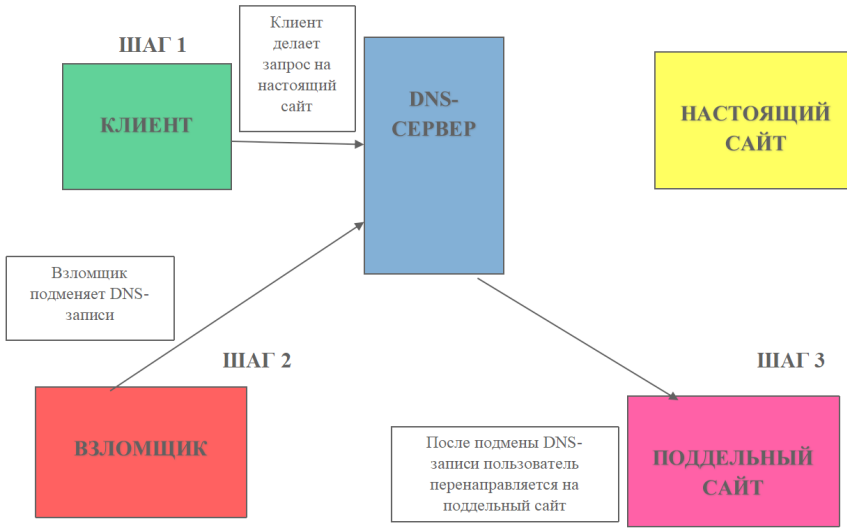

Рис. 1. Процесс реализации DNS-спуфинга

чение для кражи конфиденциальной информации.

Чтобы предотвратить данную атаку необходимо осуществлять мониторинг данных DNS это может помочь установить наличие в трафике отклоняющихся явлений, действий пользователей. Обнаружение отравления DNS-кэшем является сложным процессом, существует несколько мер безопасности, которые необходимо принять, чтобы предотвратить данную атаку. Мера, предотвращающая отравление DNS-кэша, включает в себя использование DNSSEC, отключение рекурсивных запросов и т. д.

DNSSEC (расширения безопасности системы доменных имен) – использует криптографию с открытым ключом для подписи DNS-записей для этого добавляется функция проверки и позволяют системам определять, является ли адрес настоящим или фальшивым. Это поможет проверить и аутентифицировать подлинность запросов и предотвратить подделку.

Отравление кэш-памяти DNS это атака, которую злоумышленник перенаправляет пользователей на вредоносные сайты. Управляемые взломщиками серверы могут ввести в заблуждение пользователей тем самым участники ничего не подозревая скачивают вредоносное программное обеспечение или предоставляют конфиденциальные данные. Чтобы защититься от данной угрозы, необходимо применить методы защиты и обеспечить безопасность DNS-инфраструктуры.

Еще одной из ключевых атак является атака нулевого дня. Атака нулевого дня (zero-day) — это атака, к которой никто не готов. Уязвимость нулевого дня – это ошибка в программном обеспечении, о которой не знает производитель.

Данная уязвимость нулевого дня дает атакующему преимущество во времени. Чтобы реализовать защиту от уязвимостей нулевого дня необходима многоуровневая система защиты, которая включает в себя регулярное обновление программного обеспечения, внедрение систем обнаружения и предотвращения вторжений. Для того, чтобы снизить риск необходимо придерживаться следующих правил.

Схема демонстрирует этапы по предотвращению атаки нулевого дня. Она показывает какие этапы необходимы, чтобы организовать комплексную защиту системы, выявлять уязвимости на ранних этапах и быстро реагировать на угрозы. Каждый ее компонент необходим для создания многоуровневой системы безопасности, которая минимизирует риски эксплуатации нулевых уязвимостей.

Уязвимость нулевого дня – это проверка процессов и архитектуры. При недостаточном анализе процессов, архитектуры и контроля за безопасностью подвергается риску вся инфраструктура. Чтобы минимизировать риск необходимо выполнять следующие этапы - регулярно проводить аудит архитектуры, внедрять многоуровневую защиту, проводить пентесты на уязвимые места и подготовить систему к быстрому реагированию при обнаружении уязвимостей.

Рассмотрим пример атаки на внутренние сети с помощью компрометированных IoT-устройств – это взлом устройств интернета вещей (камеры видеонаблюдения, термоста-

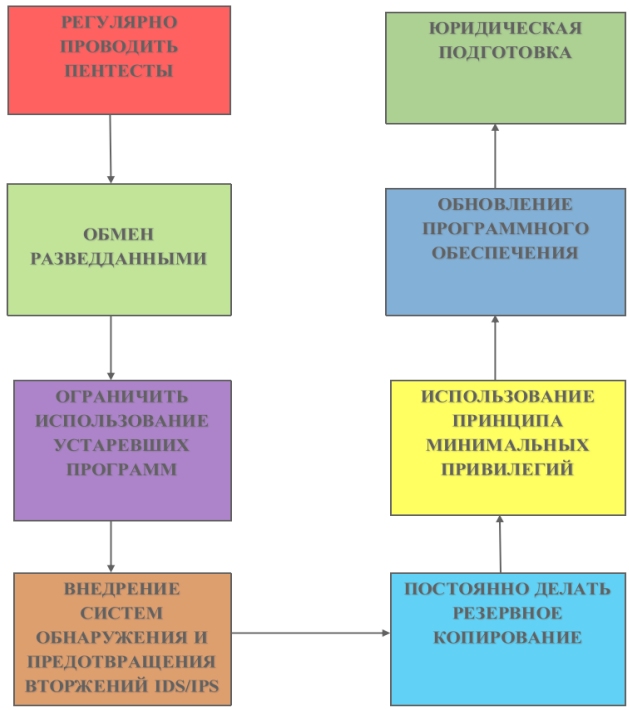

Рис. 2. Этапы, необходимые для минимизации атаки нулевого дня

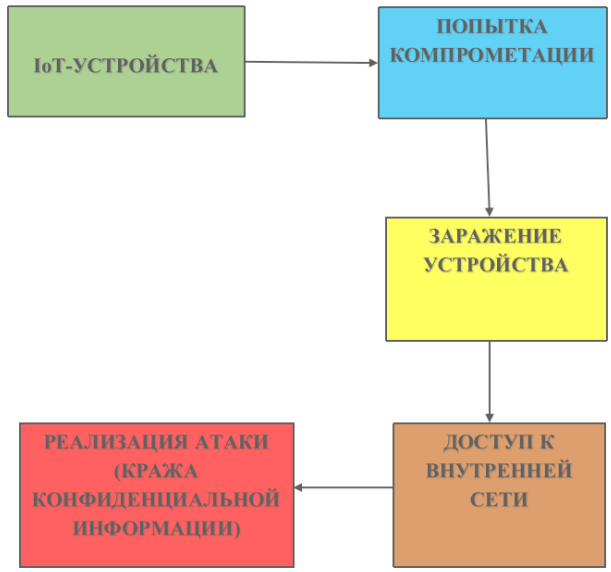

Рис. 3. Этапы атаки с помощью компрометированных IoT-устройств.

ты и другие гаджеты) для получения доступа к корпоративной сети.

Схема предназначена для визуализации процесса угрозы, связанной с использованием IoT-устройств в внутренней сети организации. Она помогает понять, как уязвимости в IoT-устройствах могут привести к серьезным последствиям начиная от взлома и формирования «ботнетов» до получения доступа злоумышленников к внутренним системам.

Данная схема демонстрирует меры защиты сети для обеспечения безопасности IoT-устройств, для организации многоуровневой защиты инфраструктуры и внутренней сети. Она показывает последовательность и взаимосвязь ключевых мер, направленных на снижение рисков и предотвращение атак.

Таким образом, правильная конфигурация каждой из представленных мер позволяют сформировать надежную защиту сети, снизить вероятность успешных атак и быстро реагировать на инциденты. Это обеспечивает

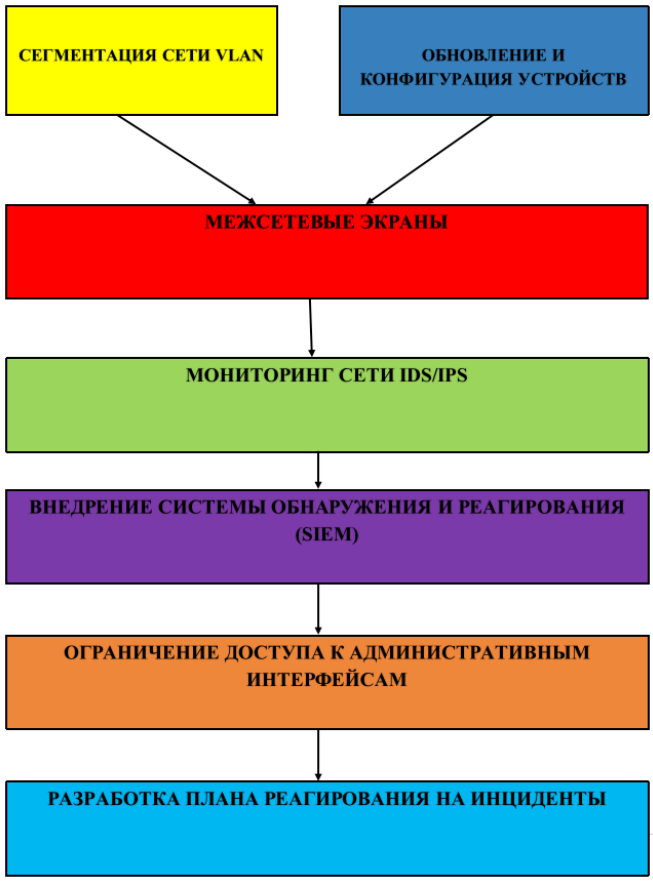

Рис. 4. Меры защиты сети

Таблица 1

Рекомендации для повышения уровня защиты информационной инфраструктуры. План мер по информационной безопасности

РЕКОМЕНДАЦИИ ПО ПОВЫШЕНИЮ УРОВНЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ	ОПИСАНИЕ	ЛИЦА, ОТВЕЧАЮЩИЕ ЗА ВЫПОЛНЕНИЕ РЕКОМЕНДАЦИЙ	ПЕРИОДИЧНОСТЬ ВЫПОЛНЕНИЯ
Внедрение многоуровневой системы защиты	Использование межсетевых экранов, VPN и т. д.	Администраторы сети, инженеры по информационной безопасности	Постоянно, при любых изменениях
Обновление программного обеспечения	Обновление ОС, приложений для минимизации уязвимостей	Системные администраторы и т. д.	Ежемесячно, как только будет совершенно обновление
Проводить пентесты и аудит безопасности	Тестирование уязвимостей	Внутренние специалисты по аудиту, которые помогают устранять внутренние риски, улучшать процессы в соответствии стандартам. Внешние специалисты по аудиту – независимая проверка они обеспечивают независимость и объективность оценки	Раз в полгода

Продол. табл. 1

Разработка и тестирование планов реагирования	Отработка сценариев реагирования на инциденты	Руководитель по ИБ, команда для быстрого реагирования на происшествие	Раз в год после того, как произойдет обновление политики внутреннего контроля и аудита

стабильную работу информационной системы, защищает конфиденциальность данных и способствует выполнению требований безопасности в организации. В результате создается многоуровневая система защиты, которая адаптирована к современным угрозам и обеспечивает высокий уровень безопасности всей инфраструктуры.

Рассмотренные в статье угрозы являются совокупностью атак наиболее опасные и тяжело обнаруживаемые. Для повышения уровня кибербезопасности необходимо разработать и внедрить комплексную политику безопасности, которая включает автоматизацию обновлений программного обеспечения, регулярные тесты и аудит систем, обучение сотрудников, создание планов реагирования на инциденты и постоянный мониторинг инфраструктуры. Внедрение многоуровневых систем защиты поможет снизить риски и обеспечить надежную работу информационной системы, защиту данных и соблюдение требований безопасности.

Рекомендации предназначены для планирования мероприятий, направленных на повышение уровня информационной безопасности организации, где приведены ключевые рекомендации для обеспечения четкого распределения ответственности, контроля сроков и реализации данных мер.

Таким образом, создание и регулярное обновление плана мер по информационной безопасности является ключевым аспектом для комплексной защиты информации организации. Он помогает обеспечить решительный подход к минимизации рисков и повысить уровень защиты данных для обеспечения безопасной работы информационных систем.

Заключение

В результате проведенного исследования были рассмотрены методы и средства мониторинга обнаружения и противодействия атак:

1. Атаки DNS spoofing.

2. Уязвимости нулевого дня.

3. Применение IoT-устройств злоумышленниками.

В результате исследования, были выполнены следующие задачи:

1) разработан комплексный подход к повышению уровня защиты, включающий внедрение многоуровневых систем безопасности;

2) Рассмотрены методы использования криптографических технологий для организации регулярных тестов и аудитов;

3) создан безопасный план реагирования на инциденты.

Таким образом, данный комплексный подход к обеспечению информационной безопасности и предложенные практические рекомендации по созданию надежных систем защиты, что является актуальным и необходимым для корпоративных организаций в настоящее время.

Литература

1. Программно-аппаратные средства защиты информации [Электронный ресурс]: учебное пособие: в 3-х ч. / В. А. Гриднев, Ю. А. Губсков, А. С. Дерябин, А. В. Яковлев. – Тамбов: Издательский центр ФГБОУ ВО «ТГТУ» 2023.

2. Информационная безопасность: учеб. пособие / авт.-сост. И74 И. Б. Тесленко [и др.] / под общ. ред. проф. И. Б. Тесленко; Владим. гос. ун-т им. А. Г и Н. Г. Столетовых. – Владимир: Изд-во ВлГУ, 2023 – 212 с. – ISBN 978-5-9984-1783-2.

3. ГОСТ Р 50922-96. Защита информации. Основные требования и определения.

4. ГОСТ Р 51188-98. Защита информации. Испытания программных средств на наличие компьютерных вирусов. Типовое руководство.

5. ГОСТ Р 51275-99. Объекты информатизации. Факторы, воздействующие на информацию.

6. ГОСТ Р ИСО 7498-2-99. Информационная технология. Взаимосвязь открытых систем базовая эталонная модель. Часть 2. Архитектура защиты информации.

7. ГОСТ Р 51624-2000. Защита информации. Автоматизированные системы в защищенном исполнении. Общие положения.

8. ГОСТ Р ИСО/МЭК 15408-2002. Информационные технологии. Методы и средства обеспечения безопасности. Критерии безопасности информационных технологий.

9. Келдыш, Наталья Всеволодовна К 34 Системная защита информации компьютерных сетей. Учебное пособие – М.: Мир науки, 2022 – Сетевое издание. Режим доступа: https://izd-mn.com/PDF/43MNNPU22.pdf – Загл. с экрана.

10. Закон РФ «Об информации, информатизации и защите информации».

11. Шаньгин В.Ф. Комплексная защита информации в корпоративных системах: учеб. пособие / В.Ф.Шаньгин. – М.: ИД «ФОРУМ»: ИНФРА-М, 2022 – 592 с.

12. Международный стандарт ITU-T Rec. X.816 – «Модель оценки уровня безопасности».

13. Международный стандарт ITU-T Rec. X.813 – «Модель управления безопасностью».

14. Международный стандарт ITU-T Rec. X.800 «Рекомендуемые практики по обеспечению безопасности сетей».

15. Международный стандарт ITU-T Rec. X.814 – «Модель защиты сети».

16. Актуальные вопросы выявления сетевых атак. Александр Астахов. CISA. Информационный бюллетень Jet Info, № 3 (106) – Москва, 2022.

References

1. Programmno-apparatnyye sredstva zashchity informatsii [Elektronnyy resurs]: uchebnoye posobiye: v 3-kh ch. / V. A. Gridnev, YU. A. Gubskov, A. S. Deryabin, A. V. Yakovlev. – Tambov: Izdatel'skiy tsentr FGBOU VO «TGTU» 2023.

2. Informatsionnaya bezopasnost': ucheb. posobiye / avt-sost. I74 I. B. Teslenko [i dr.] / pod obshch. red. prof. I. B. Teslenko; Vladim. gos. un-t im. A. G i N. G. Stoletovykh. – Vladimir: Izd-vo VlGU, 2023 – 212 s. – ISBN 978-5-9984-1783-2.

3. GOST R 50922-96. Zashchita informatsii. Osnovnyye trebovaniya i opredeleniya. 4. GOST R 51188-98. Zashchita informatsii. Ispytaniya programmnykh sredstv na nalichiye komp'yuternykh virusov. Tipovoye rukovodstvo.

5. GOST R 51275-99. Ob"yekty informatizatsii. Faktory, vozdeystvuyushchiye na informatsiyu.

6. GOST R ISO 7498-2-99. Informatsionnaya tekhnologiya. Vzaimosvyaz' otkrytykh sistem bazovaya etalonnaya model'. Chast' 2. Arkhitektura zashchity informatsii.

7. GOST R 51624-2000. Zashchita informatsii. Avtomatizirovannyye sistemy v zashchishchennom ispolnenii. Obshchiye polozheniya.

8. GOST R ISO/MEK 15408-2002. Informatsionnyye tekhnologii. Metody i sredstva obespecheniya bezopasnosti. Kriterii bezopasnosti informatsionnykh tekhnologiy.

9. Keldysh, Natal'ya Vsevolodovna K 34 Sistemnaya zashchita informatsii komp'yuternykh setey. Uchebnoye posobiye – M.: Mir nauki, 2022 – Setevoye izdaniye. Rezhim dostupa: https://izd-mn.com/PDF/43MNNPU22.pdf – Zagl. s ekrana.

10. Zakon RF «Ob informatsii, informatizatsii i zashchite informatsii».

11. Shan'gin V.F. Kompleksnaya zashchita informatsii v korporativnykh sistemakh: ucheb. posobiye / V.F.Shan'gin. – M.: ID «FORUM»: INFRA-M, 2022 – 592 s.

12. Mezhdunarodnyy standart ITU-T Rec. X.816 – "Model' otsenki urovnya bezopasnosti".

13. Mezhdunarodnyy standart ITU-T Rec. X.813 – "Model' upravleniya bezopasnost'yu".

14. Mezhdunarodnyy standart ITU-T Rec. X.800 "Rekomenduyemyye praktiki po obespecheniyu bezopasnosti setey".

15. Mezhdunarodnyy standart ITU-T Rec. X.814 – "Model' zashchity seti".

16. Aktual'nyye voprosy vyyavleniya setevykh atak. Aleksandr Astakhov. CISA. Informatsionnyy byulleten' Jet Info, № 3 (106) – Moskva, 2022.

НИКОНОВ Вячеслав Викторович, кандидат технических наук, доцент кафедры «Интеллектуальные системы информационной безопасности» Института кибербезопасности и цифровых технологий Российского технологического университета МИРЭА. 107996, г. Москва, ул. Стромынка, д. 20. E-mail: nikonov_v@mirea.ru

ТРОФИМОВ Иван Александрович, студент кафедры «Интеллектуальные системы информационной безопасности» Института кибербезопасности и цифровых технологий Российского технологического университета МИРЭА. 107996, г. Москва, ул. Стромынка, д. 20. E-mail: vanyusha_trofimov_01@bk.ru

NIKONOV Vyacheslav Viktorovich, Candidate of Technical Sciences, Associate Professor, "Intelligent Information Security Systems," Institute of Cybersecurity and Digital Technologies, MIREA Russian Technological University. 107996, Moscow, Stromynka Street, 20. Email: nikonov_v@mirea.ru

TROFIMOV Ivan Alexandrovich, student, "Intelligent Information Security Systems," Institute of Cybersecurity and Digital Technologies, MIREA Russian Technological University. 107996, Moscow, Stromynka Street, 20. E-mail: vanyusha_trofimov_01@bk

УДК 004.852 Вестник УрФО № 2(60) / 2026, с. 76–87

Частикова В.А., Марченко В.А.

DOI: 10.14529/secur260208

СОСТЯЗАТЕЛЬНАЯ РОБАСТНОСТЬ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ: СИСТЕМАТИЧЕСКИЙ ОБЗОР УНИВЕРСАЛЬНЫХ АТАК, МЕТОДОВ ЗАЩИТЫ И ТРЕБОВАНИЙ К ИХ АУДИТУ

В работе обобщены результаты исследований состязательной (adversarial) робастности моделей глубокого обучения. Обзор подготовлен по протоколу PRISMA 2020: из 540 первоначально найденных публикаций формальным критериям включения удовлетворили 32 источника. Основное внимание уделено универсальным состязательным возмущениям (UAP) и способам защиты от них. Авторы прослеживают происхождение явления, приводят таксономию атак и разбирают основные алгоритмы построения UAP, после чего сопоставляют между собой современные защиты – состязательное обучение, TRADES, случайное сглаживание – и эталонные бенчмарки AutoAttack, RobustBench и ImageNet-C. Отдельный раздел соотносит международные результаты с действующей в России нормативной базой: приказами ФСТЭК России № 17, № 21 и № 239, ФЗ-187, ГОСТ Р ИСО/МЭК 27001-2021 и ГОСТ Р 59276-2020. Как показывает анализ, универсальность возмущений вместе с их переносимостью между архитектурами делает атаки практически значимыми, тогда как формальные гарантии среди известных защит дают только методы сертифицированной робастности. В заключение выделены вопросы, которые в национальной методической базе требуют первоочередной проработки: физически реализуемые UAP, защита крупных мультимодальных моделей и стандартизация процедур аудита значимых объектов критической информационной инфраструктуры по устойчивости к adversarial-воздействиям.

Ключевые слова: *глубокое обучение, состязательные атаки, универсальные состязательные возмущения, состязательное обучение, случайное сглаживание, переносимость атак, AutoAttack, RobustBench, информационная безопасность, ФСТЭК России.*

ADVERSARIAL ROBUSTNESS OF MACHINE LEARNING MODELS: A SYSTEMATIC REVIEW OF UNIVERSAL ATTACKS, DEFENSE METHODS, AND AUDIT REQUIREMENTS

The paper presents a systematic review of adversarial robustness of deep learning models, conducted per the PRISMA 2020 protocol. Of 540 initially identified publications, 32 sources passed the formal inclusion and exclusion criteria. The focus is on Universal Adversarial Perturbations (UAP) and defense approaches. The origins and taxonomy of attacks are systematized; key UAP generation algorithms are reviewed; modern defense methods (adversarial training, TRADES, randomized smoothing) and benchmarks (AutoAttack, RobustBench, ImageNet-C) are analyzed. A separate section maps international results onto the Russian regulatory framework – FSTEC orders No. 17, No. 21, No. 239, Federal Law 187-FZ, GOST R ISO/IEC 27001-2021 and GOST R 59276-2020. Universality and cross-architecture transferability make these attacks practically significant; among existing defenses, only certified methods provide formal guarantees. Priority directions for the national methodological framework are formulated: physically realizable UAPs, defense of large multimodal models, and standardization of adversarial-robustness audit procedures for significant objects of critical information infrastructure.

Keywords: *deep learning, adversarial attacks, universal adversarial perturbations, adversarial training, randomized smoothing, transferability, AutoAttack, RobustBench, information security, FSTEC of Russia.*

Введение

Сегодня глубокое обучение применяется в компьютерном зрении, обработке естественного языка и системах поддержки принятия решений, в том числе в таких ответственных областях, как медицинская диагностика, автономный транспорт, биометрия и средства информационной безопасности [1]. За последнее десятилетие, однако, у нейросетевых моделей обнаружился новый класс уязвимостей – состязательные примеры (adversarial examples). Это специально подобранные входные данные, которые отличаются от исходных небольшими, незаметными для человека возмущениями, но при этом способны полностью изменить решение классификатора [2, 3]. Число работ по этой теме идет на сотни, поэтому обобщение накопленных результатов само по себе превратилось в отдельную научную задачу [4, 5].

Среди таких атак выделяются универсальные состязательные возмущения (UAP) – единый шумовой паттерн, не привязанный к конкретному входу: будучи наложенным на произвольное изображение из распределения данных, он с высокой вероятностью вызывает у модели ошибку [6]. От классических состязательных атак UAP отличаются масштабируемостью. Возмущение достаточно вычислить один раз, после чего его можно применять многократно и даже физически – например, напечатав на наклейке или билборде либо нанеся на одежду [7]. Из-за этого UAP представляют не только академический интерес, но и реальную угрозу для биометрического контроля доступа, систем видеоанали-

тики, автономного транспорта и медицинской диагностики [8].

Отечественная нормативная база учитывает эту угрозу лишь частично. Приказы ФСТЭК России № 17, № 21 и № 239 [23, 24, 25], Федеральный закон № 187-ФЗ [26], ГОСТ Р ИСО/МЭК 27001-2021 [27] и ГОСТ Р 59276-2020 [28] задают общие требования к доверию и защите, однако в них нет ни прикладных методик оценки устойчивости моделей машинного обучения к UAP, ни пороговых значений метрик, ни обязательного перечня атакующих алгоритмов для аттестационного тестирования. Тем самым между международной исследовательской практикой и российской системой сертификации значимых объектов критической информационной инфраструктуры (КИИ) образуется методологический разрыв.

Цель настоящего обзора – привести в систему ключевые публикации 2013–2025 гг. по двум связанным направлениям: с одной стороны, природа и алгоритмы генерации универсальных состязательных возмущений, с другой – подходы к защите моделей, как эмпирические, так и сертифицированные. Попутно мы соотносим эти результаты с требованиями ФСТЭК России и национальных стандартов и указываем направления, которые в нормативно-методической базе следует проработать в первую очередь.

1. Методология обзора

Работа выполнена по рекомендациям PRISMA 2020 [29]. Литературу мы искали с января по апрель 2026 г. в базах Google Scholar, IEEE Xplore, ACM Digital Library, OpenReview, arXiv и Scopus. Запросы строились по комбинациям ключевых терминов: «adversarial examples», «universal adversarial perturbation», «adversarial robustness», «certified robustness», «adversarial training», «adversarial attack defense». Поиск охватывал период 2013–2025 гг.; нижняя граница задана работой Szegedy и др. [2]. В качестве метаисточников использованы RobustBench [13] и обзоры [4, 5]. Этапы фильтрации сведены в табл. 1.

Таблица 1

Этапы отбора публикаций (PRISMA 2020 flow)

Этап отбора	Количество записей
Identification (всего найдено)	540
Удалено дубликатов	89
Screening по заголовку и аннотации	451
Исключено по тематике / нерелевантности	241
Eligibility (оценка полного текста)	210
Исключено по критериям невключения	178
Included (вошло в обзор)	32

Критерии включения. К рассмотрению допускались работы: опубликованные в реферируемом издании или на конференциях ранга A* (NeurIPS, ICML, ICLR, CVPR, ICCV, ECCV, IEEE S&P, USENIX Security, ACM CCS, NDSS); содержащие экспериментальную либо теоретическую новизну в области состязательной робастности; имеющие открытую реализацию или включенные в RobustBench; индексируемые с DOI или arXiv-идентификатором.

Критерии исключения. Из дальнейшей обработки исключались:

1) препринты без последующей рецензируемой публикации со сроком более 24 месяцев;

2) работы по adversarial-атакам в неинформационно-визуальных доменах вне общеметодологических вопросов;

3) работы без воспроизводимых результатов (нет кода, гиперпараметров или данных);

4) дубликаты публикаций (расширенные журнальные версии учитывались как один источник);

5) тезисы и материалы воркшопов без полного рецензирования;

6) источники старше 2013 г., кроме классических работ, заложивших методологический базис;

7) работы, не прошедшие минимальную оценку качества по адаптированной шкале.

Качество каждой включенной работы оценивалось по четырем критериям: воспроизводимость (открытый код и веса), статистическая обоснованность (доверительные ин-

тервалы или тесты значимости), независимость данных (использование стандартных бенчмарков) и – при заявлении новых защитных механизмов – проверка под адаптивной атакой [14]. На публикации 2021–2025 гг. пришлось 34 % включенных источников.

2. Результаты

2.1. Истоки проблемы и таксономия атак

Впервые состязательные примеры описали Szegedy и др. в 2013 г., изучая «странные» свойства глубоких нейронных сетей [2]. Они обнаружили, что почти для каждого правильно классифицируемого изображения можно построить визуально неотличимый от него вариант, который сеть отнесет к заранее выбранному ошибочному классу. Двумя годами позже Goodfellow и др. объяснили этот эффект линейностью сетей в пространствах высокой размерности и предложили быстрый способ генерации таких примеров – Fast Gradient Sign Method (FGSM) [3].

Атаки принято классифицировать сразу по нескольким независимым осям: по уровню доступа к модели (white-box, black-box, серобоксный), по цели (нецелевые и целевые), по этапу жизненного цикла (отравление обучающих данных либо атаки на этапе вывода, evasion), по используемой метрике возмущения (L_∞, L_2, L_0), а также по специфичности – индивидуальные или универсальные [4, 5, 9]. В систематических обзорах такое деление применяется как стандартное [4].

Наряду с эмпирическими методами велась и теоретическая оценка робастности. Carlini и Wagner установили, что многие защиты 2016–2017 гг. держались на «маскировании градиента» и легко пробивались сильной итеративной атакой [7]. Athalye и др. распространили этот вывод шире: из девяти защит, представленных на ICLR 2018, семь удалось обойти, скомпенсировав обнаруженную маскировку [14]. Так сформировался методологический стандарт: каждую новую защиту следует проверять адаптивной атакой, построенной именно против нее.

2.2. Универсальные состязательные возмущения

Понятие универсального состязательного возмущения предложили Moosavi-Dezfooli и др. [6]. Для произвольной обученной модели f и распределения данных μ они построили единый вектор v с малой L_∞-нормой, при наложении которого более 80 % изображений из μ классифицируются неверно. Показательно, что возмущение, рассчитанное на VGG-19, успешно атакует и GoogLeNet, и ResNet-152: это говорит об общих геометрических уязвимостях, присущих обученным сетям [6, 9].

Геометрически существование UAP объясняют так: в высокоразмерном пространстве признаков границы принятия решения проходят систематически близко к естественному распределению данных – вдоль некоторого «направления наименьшей устойчивости» [6]. Именно эта общая для разных входов уязвимость и позволяет одному вектору обманывать модель на широком классе изображений. Переносимость UAP, в свою очередь, зависит от архитектурного сходства моделей: внутри семейства сверточных сетей возмущения переносятся хорошо, тогда как между CNN и трансформерами перенос заметно слабее [9, 30].

На практике опасность UAP усиливается тем, что их можно реализовать физически. Eykholt и др. с помощью наклеек успешно атаковали классификаторы дорожных знаков: в полевых испытаниях знак «STOP» принимался за «Speed Limit 45» в 84 % случаев – при разных ракурсах съемки и освещении [10]. Для систем автономного транспорта и видеоаналитики на значимых объектах КИИ – это прямая угроза [25].

2.3. Эволюция алгоритмов генерации UAP

В исходном алгоритме Moosavi-Dezfooli универсальное возмущение набирается итеративно – как накопление минимальных индивидуальных возмущений (DeepFool) с проекцией в шар L_∞-нормы радиуса ε [6]. Алгоритм несложен, однако требует обучающих изображений и сходится лишь за сотни итераций. Дальнейшие работы развивались в двух направлениях – ускорение сходимости и снижение требований к данным.

Первое направление – генеративные подходы. В методе NAG Mopuri и др. обучают условную генеративную сеть, которая сама порождает UAP, оптимизируя функцию обмана модели-жертвы [11]. По сравнению с итеративным методом Moosavi-Dezfooli при той же норме ε это дает более разнообразные возмущения и до 5 % более высокий fooling rate. Та же идея заложена в GAP, где к состязательной постановке задачи добавлены возможности GAN [16].

Второе направление – data-free-атаки, которым доступ к обучающему набору вообще не нужен. Вместо обучающих данных они

опираются на априорные сведения о статистике активаций внутренних слоев сети [11]. На практике это особенно опасно: злоумышленнику не требуется представительная выборка из домена жертвы, а fooling rate при ε = 10/255 удерживается на уровне 60–70 % – почти как у методов, использующих данные.

2.4. Подходы к защите

Защиты обычно делят на три группы: состязательное обучение (AT), то есть обучение на смеси чистых и состязательных примеров; методы предобработки и сглаживания входа; наконец, сертифицированные методы. Наиболее сильной эмпирической защитой остается состязательное обучение в формулировке Madry и др. [12]: модель решает min-max задачу, минимизируя потери по параметрам и одновременно максимизируя их по PGD-возмущениям в шаре радиуса ε. Защиты сопоставлены в табл. 2.

Таблица 2

Сравнение основных подходов к защите от состязательных атак

Метод	Тип гарантий	Снижение FR (UAP)	Потеря Top-1, п.п.	Стоимость обучения/инференса	Источник
AT (Madry, PGD-AT)	эмпирические	−45..−55 п.п.	10–14	обучение в 5–10× медленнее	[12]
TRADES	эмпирические	−40..−50 п.п.	8–11	обучение в 5–8× медленнее	[18]
Free AT (Shafahi)	эмпирические	−35..−45 п.п.	12–15	обучение в ~1.5× медленнее	[19]
Fast AT (Wong)	эмпирические	−30..−40 п.п.	10–13	обучение в ~2× медленнее	[17]
Gowal AT (доп. данные)	эмпирические	−50..−58 п.п.	5–8	обучение в ~6× медленнее	[31]
Distillation (Papernot)	эмпирические, обходимо	малое	≤2	обучение в 2× медленнее	[15]
Randomized Smoothing	сертифицированные	−50..−55 п.п.	8–12	инференс в 50–100× медленнее	[20]

TRADES [18] выносит баланс между чистой и робастной точностью в явный гиперпараметр β и на ImageNet и CIFAR-10 устойчиво обходит классическое AT по метрике AutoAttack. Free AT [19] и Fast AT [17] решают другую задачу – удешевляют обучение ценой небольшого снижения робастности. Наконец, Gowal и др. [31] и Rebuffi и др. [32] показали, насколько сильно AT выигрывает от дополнительных или синтезированных данных: на CIFAR-10 при ε = 8/255 робастная точность поднимается с 53 % у базовой модели Madry до 66 %.

Сертифицированные методы устроены принципиально иначе. Случайное сглаживание Cohen, Rosenfeld и Kolter [20] дает формальную гарантию: для входа x при заданном σ вычисляется радиус r, внутри которого ответ классификатора заведомо не меняется. Гарантия носит вероятностный характер, но при достаточном числе шумовых проходов превращается в практически применимую сертификацию. Цена этого – рост времени инференса в 50–100 раз, из-за чего метод плохо подходит для систем реального времени.

Отдельного внимания заслуживает проблема мнимой защиты. Дистилляцию, предложенную Papernot и др. [15], вскоре обошли Carlini и Wagner своей C&W-атакой [7]. Точно так же простые преобразования входа – квантование, фильтрация, JPEG-сжатие – пробиваются адаптивными атаками [14]. Отсюда и общее методологическое требование: защиту следует проверять не только под FGSM или PGD, но и под адаптивным противником.

2.5. Бенчмарки и стандартизированные оценки

Невоспроизводимость результатов и их сильная зависимость от выбранного атакующего алгоритма долго оставались системной проблемой области. Чтобы снять эту зависимость, Croce и Hein предложили AutoAttack [21] – ансамбль из четырех атак (APGD-CE, APGD-DLR, FAB и Square) без свободных гиперпараметров; сегодня он де-факто служит эта-

лоном при оценке робастности. На AutoAttack опирается и лидерборд RobustBench [13], где собрано свыше 100 моделей с проверенными значениями стандартной и робастной точности.

Устойчивость к естественным искажениям измеряют отдельным бенчмарком – ImageNet-C / CIFAR-10-C [22], в котором предусмотрено 19 типов искажений на пяти уровнях интенсивности. Состязательную робастность он напрямую не охватывает, зато позволяет проверить, насколько связаны обычная и адверсарная устойчивость. Связь оказалась слабой, что лишний раз подтверждает: защита от UAP – самостоятельная задача. По данным Bai и др. [30] и Sehwag и др. [33], при одинаковых условиях обучения трансформеры (ViT, Swin) робастнее сверточных сетей, однако под адаптивными атаками это преимущество в значительной степени нивелируется.

3. Обсуждение

3.1. Сводный анализ утверждений

Из всей рассмотренной литературы можно вычленить несколько устойчивых утверждений, за которыми стоит разная по силе доказательная база. В табл. 3 они сведены в формате «утверждение – оценка силы доказательств – источники».

Табл. 3 показывает, что к 2025 году по ряду фундаментальных вопросов в сообществе сложился консенсус: существование UAP, их переносимость внутри семейства CNN, превосходство AT над тривиальными защитами, польза дополнительных данных. Доказательства средней силы остаются там, где сравнивают трансформеры и CNN, оценивают защиты для новых архитектур (Swin, ConvNeXt) или обсуждают применимость сертифицированных методов в продакшене. Заметный методологический сдвиг произошел в 2018–2021 гг.: после работы Athalye и др. [14] и появления AutoAttack [21] оценка под адаптивным противником и публи-

Таблица 3

Ключевые утверждения о состязательной робастности и сила доказательной базы

Утверждение	Сила	Обоснование	Источники
Глубокие сети уязвимы к состязательным примерам	Strong (10/10)	Воспроизведено сотнями работ, теоретическое объяснение через локальную линейность	[2, 3]
Существуют UAP с FR ≥ 80 % при ε = 10/255	Strong (9/10)	Подтверждено для 6+ архитектур на ImageNet	[6, 9, 11]
UAP переносимы между CNN-архитектурами	Strong (8/10)	Подтверждено для VGG, ResNet, DenseNet, GoogLeNet	[6, 9]
UAP слабо переносятся между CNN и трансформерами	Moderate (7/10)	Несколько независимых наблюдений на ViT и Swin	[9, 30]
Состязательное обучение Madry — сильнейший эмпирический базис	Strong (9/10)	Лидер RobustBench 2018–2021, независимо воспроизведено	[12, 13, 21]
Доп. / синтетические данные существенно повышают робастность AT	Strong (8/10)	CIFAR-10: рост робастной точности с 53 % до 66 %	[31, 32]
TRADES обеспечивает лучший компромисс точности/робастности	Moderate (7/10)	Подтверждено на CIFAR-10/100, частично на ImageNet	[18]
Случайное сглаживание — единственный масштабируемый сертифицированный метод	Strong (8/10)	Вероятностные гарантии, доказанные теоремами	[20]
Простые защиты (дистилляция, фильтрация) обходятся адаптивной атакой	Strong (9/10)	Систематическая проверка 9 защит ICLR 2018	[7, 14]
Физически реализуемые атаки эффективны в реальных условиях	Moderate (7/10)	Несколько демонстраций на знаках/одежде	[10]
Состязательная робастность требует увеличенного объема данных	Moderate (6/10)	Теоретический анализ для линейных моделей	[14]

кация моделей в RobustBench стали нормой. При этом с 2020 г. лидеры лидерборда прибавляют на CIFAR-10 лишь доли процента робастной точности в год, и каждый такой шаг обходится во все большие вычисления [13, 31].

3.2. Применимость в контексте российской нормативной базы

В России защита значимых информационных систем регулируется актами ФСТЭК и национальными стандартами. Приказ № 17 [23] относится к государственным ИС, приказ № 21 [24] – к персональным данным, а приказ № 239 [25] во исполнение ФЗ-187 [26] – к значимым объектам КИИ. Общие требования к системе менеджмента ИБ задает ГОСТ Р ИСО/МЭК 27001-2021 [27], тогда как доверие к системам ИИ прямо регулирует единственный национальный стандарт – ГОСТ Р 59276-2020 [28]. На международном уровне эти вопросы затрагивают ИСО/МЭК 23894:2023 [34] и NIST AI RMF 1.0 [35].

Сопоставление этих документов с проблематикой состязательной робастности приведено в табл. 4.

Таблица 4

Покрытие угрозы состязательных атак нормативными документами

Документ	Объект регулирования	Покрытие adversarial-угроз
ГОСТ Р ИСО/МЭК 27001-2021 [27]	СМИБ организации	Отсутствует прямое регулирование
ГОСТ Р 59276-2020 [28]	Доверие к системам ИИ	Упомянуто, без методик оценки
Приказ ФСТЭК № 17 [23]	Государственные ИС	Косвенно через тестирование защищенности
Приказ ФСТЭК № 21 [24]	ИС персональных данных	Косвенно через анализ уязвимостей
Приказ ФСТЭК № 239 [25]	Значимые объекты КИИ	Косвенно через периодический аудит
ФЗ-187 [26]	Безопасность КИИ	Общие требования, без специфики ИИ
БДУ ФСТЭК [36]	Банк данных угроз	Категория угроз для ИИ формируется
ИСО/МЭК 23894:2023 [34]	Управление рисками ИИ	Общие положения по adversarial-рискам
NIST AI RMF 1.0 [35]	Управление рисками ИИ	Adversarial-устойчивость как часть Trust

Из табл. 4 видно главное: ни один из действующих российских документов не содержит формализованной методики оценки устойчивости моделей машинного обучения к UAP. Нигде не закреплены ни пороговые значения fooling rate, ни обязательный перечень атакующих алгоритмов для аттестации, ни требование публиковать воспроизводимые результаты аудита. В БДУ ФСТЭК [36] категория угроз для систем ИИ пока только формируется, и специфика adversarial-атак – UAP, отравление данных, инверсия модели – представлена лишь в общих чертах. О том же пробеле пишут Намиот и Ильюшин [37].

На практике это оборачивается разрывом между международными исследовательскими стандартами – AutoAttack, RobustBench, PRISMA – и отечественной системой аттестации. Разработчикам защищенных систем компьютерного зрения (биометрия, видеоаналитика на объектах КИИ, медицинская диагностика) приходится либо самостоятельно переносить к себе зарубежные методики, либо формально выполнять общие требования, не проверяя систему на специфические угрозы. Чтобы устранить этот разрыв, нужен отдельный методический документ ФСТЭК, ориентированный именно на adversarial-угрозы.

3.3. Матрица исследовательских пробелов

Табл. 5 сводит открытые вопросы и типы защитных методов в единую матрицу пробе-

Таблица 5

Матрица исследовательских пробелов

Открытый вопрос / Метод	AT/PGD	TRADES	Random. Smoothing	Detection
Робастность для трансформеров (ViT, Swin)	3	1	GAP	GAP
Защита от физически реализуемых UAP	2	GAP	GAP	GAP
Применимость к мультимодальным моделям (CLIP)	GAP	GAP	GAP	GAP
Стандартизация в нормативных документах ФСТЭК	GAP	GAP	GAP	GAP
Гармонизация с ГОСТ Р 59276-2020	GAP	GAP	GAP	GAP
Сертифицированный аудит больших LLM	GAP	GAP	1	GAP

лов: пометка GAP означает, что систематических исследований по соответствующей ячейке нет, а число показывает, сколько публикаций по ней удалось найти.

Заключение

К настоящему времени состязательная робастность оформилась в самостоятельную область со своей таксономией атак, набором эталонных алгоритмов и воспроизводимыми бенчмарками. Центральное место в ней занимают универсальные состязательные возмущения: сочетая свойства классических атак с масштабируемостью и переносимостью между моделями, они становятся одной из наиболее реалистичных угроз для развернутых систем компьютерного зрения, в том числе на значимых объектах КИИ.

Защиты выстраиваются в довольно ясную иерархию. На одном полюсе – эмпирические методы во главе с состязательным обучением Madry и его ускоренными вариантами (Free AT, Fast AT, Gowal AT) и компромиссным TRADES; на другом – сертифицированные методы, прежде всего случайное сглаживание. Ни один из классов не свободен от ограничений: эмпирические защиты не гарантируют устойчивости и со временем обходятся адаптивной атакой, а сертифицированные дороги в инференсе и плохо масштабируются на сложные задачи.

Сопоставление международной практики с российской нормативной базой обнажает методологический пробел: ни ГОСТ Р ИСО/МЭК 27001-2021, ни ГОСТ Р 59276-2020, ни приказы ФСТЭК № 17, № 21 и № 239 не предъявляют формализованных требований к оценке устойчивости моделей машинного обучения к UAP. Открытыми мы считаем четыре направления: (1) оценка робастности новых архитектур – трансформеров, диффузионных и мультимодальных моделей; (2) защита от физически реализуемых UAP, важная для биометрии и автономного транспорта; (3) подготовка методических документов ФСТЭК по аттестационному тестированию ИИ-компонентов значимых объектов КИИ; (4) сертифицированная робастность больших языковых моделей.

Литература

1. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature. – 2015. – Vol. 521, no. 7553. – P. 436–444. DOI: 10.1038/nature14539.

2. Szegedy C., Zaremba W., Sutskever I., Bruna J., Erhan D., Goodfellow I., Fergus R. Intriguing properties of neural networks // International Conference on Learning Representations (ICLR). – 2014. – arXiv:1312.6199 [Электронный ресурс]. – URL: https://arxiv.org/abs/1312.6199 (дата обращения: 15.05.2026).

3. Goodfellow I.J., Shlens J., Szegedy C. Explaining and harnessing adversarial examples // International Conference on Learning Representations (ICLR). – 2015. – arXiv:1412.6572 [Электронный ресурс]. – URL: https://arxiv.org/abs/1412.6572 (дата обращения: 15.05.2026).

4. Biggio B., Roli F. Wild patterns: Ten years after the rise of adversarial machine learning // Pattern Recognition. – 2018. – Vol. 84. – P. 317–331. – DOI: 10.1016/j.patcog.2018.07.023.

5. Akhtar N., Mian A. Threat of adversarial attacks on deep learning in computer vision: A survey // IEEE Access. – 2018. – Vol. 6. – P. 14410–14430. – DOI: 10.1109/ACCESS.2018.2807385.

6. Moosavi-Dezfooli S.-M., Fawzi A., Fawzi O., Frossard P. Universal adversarial perturbations //

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017. – P. 1765–1773. – DOI: 10.1109/CVPR.2017.17.

7. Carlini N., Wagner D. Towards evaluating the robustness of neural networks // IEEE Symposium on Security and Privacy (S&P). – 2017. – P. 39–57. – DOI: 10.1109/SP.2017.49.

8. Finlayson S.G., Bowers J.D., Ito J., Zittrain J.L., Beam A.L., Kohane I.S. Adversarial attacks on medical machine learning // Science. – 2019. – Vol. 363, no. 6433. – P. 1287–1289. – DOI: 10.1126/science.aaw4399.

9. Tramèr F., Papernot N., Goodfellow I., Boneh D., McDaniel P. The space of transferable adversarial examples // arXiv preprint. – 2017. – arXiv:1704.03453 [Электронный ресурс]. – URL: https://arxiv.org/abs/1704.03453 (дата обращения: 15.05.2026).

10. Eykholt K., Evtimov I., Fernandes E., Li B., Rahmati A., Xiao C., Prakash A., Kohno T., Song D. Robust physical-world attacks on deep learning visual classification // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2018. – P. 1625–1634. – DOI: 10.1109/CVPR.2018.00175.

11. Mopuri K.R., Ojha U., Garg U., Babu R.V. NAG: Network for adversary generation // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2018. – P. 742–751. – DOI: 10.1109/CVPR.2018.00084.

12. Madry A., Makelov A., Schmidt L., Tsipras D., Vladu A. Towards deep learning models resistant to adversarial attacks // International Conference on Learning Representations (ICLR). – 2018. – arXiv:1706.06083 [Электронный ресурс]. – URL: https://arxiv.org/abs/1706.06083 (дата обращения: 15.05.2026).

13. Croce F., Andriushchenko M., Sehwag V., Debenedetti E., Flammarion N., Chiang M., Mittal P., Hein M. RobustBench: A standardized adversarial robustness benchmark // Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS) Datasets and Benchmarks Track. – 2021. – arXiv:2010.09670 [Электронный ресурс]. – URL: https://arxiv.org/abs/2010.09670 (дата обращения: 15.05.2026).

14. Athalye A., Carlini N., Wagner D. Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples // Proceedings of the 35th International Conference on Machine Learning (ICML). – 2018. – P. 274–283. – arXiv:1802.00420 [Электронный ресурс]. – URL: https://arxiv.org/abs/1802.00420 (дата обращения: 15.05.2026).

15. Papernot N., McDaniel P., Wu X., Jha S., Swami A. Distillation as a defense to adversarial perturbations against deep neural networks // IEEE Symposium on Security and Privacy (S&P). – 2016. – P. 582–597. – DOI: 10.1109/SP.2016.41.

16. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial nets // Advances in Neural Information Processing Systems (NeurIPS). – 2014. – Vol. 27. – arXiv:1406.2661 [Электронный ресурс]. – URL: https://arxiv.org/abs/1406.2661 (дата обращения: 15.05.2026).

17. Wong E., Rice L., Kolter J.Z. Fast is better than free: Revisiting adversarial training // International Conference on Learning Representations (ICLR). – 2020. – arXiv:2001.03994 [Электронный ресурс]. – URL: https://arxiv.org/abs/2001.03994 (дата обращения: 15.05.2026).

18. Zhang H., Yu Y., Jiao J., Xing E.P., El Ghaoui L., Jordan M.I. Theoretically principled trade-off between robustness and accuracy // Proceedings of the 36th International Conference on Machine Learning (ICML). – 2019. – P. 7472–7482. – arXiv:1901.08573 [Электронный ресурс]. – URL: https://arxiv.org/abs/1901.08573 (дата обращения: 15.05.2026).

19. Shafahi A., Najibi M., Ghiasi M.A., Xu Z., Dickerson J., Studer C., Davis L.S., Taylor G., Goldstein T. Adversarial training for free! // Advances in Neural Information Processing Systems (NeurIPS). – 2019. – Vol. 32. – arXiv:1904.12843 [Электронный ресурс]. – URL: https://arxiv.org/abs/1904.12843 (дата обращения: 15.05.2026).

20. Cohen J., Rosenfeld E., Kolter J.Z. Certified adversarial robustness via randomized smoothing // Proceedings of the 36th International Conference on Machine Learning (ICML). – 2019. – P. 1310–1320. – arXiv:1902.02918 [Электронный ресурс]. – URL: https://arxiv.org/abs/1902.02918 (дата обращения: 15.05.2026).

21. Croce F., Hein M. Reliable evaluation of adversarial robustness with an ensemble of diverse parameter-free attacks // Proceedings of the 37th International Conference on Machine Learning (ICML). – 2020. – P. 2206–2216. – arXiv:2003.01690 [Электронный ресурс]. – URL: https://arxiv.org/abs/2003.01690 (дата обращения: 15.05.2026).

22. Hendrycks D., Dietterich T. Benchmarking neural network robustness to common corruptions and perturbations // International Conference on Learning Representations (ICLR). – 2019. – arXiv:1903.12261 [Электронный ресурс]. – URL: https://arxiv.org/abs/1903.12261 (дата обращения: 15.05.2026).

23. Приказ ФСТЭК России от 11.02.2013 № 17 «Об утверждении Требований о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах» (с изм.) [Электронный ресурс] // ФСТЭК России. – URL: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-11-fevralya-2013-g-n-17 (дата обращения: 15.05.2026).

24. Приказ ФСТЭК России от 18.02.2013 № 21 «Об утверждении Состава и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных» (с изм.) [Электронный ресурс] // ФСТЭК России. – URL: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-18-fevralya-2013-g-n-21 (дата обращения: 15.05.2026).

25. Приказ ФСТЭК России от 25.12.2017 № 239 «Об утверждении Требований по обеспечению безопасности значимых объектов критической информационной инфраструктуры Российской Федерации» (с изм.) [Электронный ресурс] // ФСТЭК России. – URL: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-25-dekabrya-2017-g-n-239 (дата обращения: 15.05.2026).

26. Федеральный закон от 26.07.2017 № 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации» // Собрание законодательства Российской Федерации. – 2017. – № 31 (ч. I). – Ст. 4736.

27. ГОСТ Р ИСО/МЭК 27001-2021. Информационная безопасность, кибербезопасность и защита конфиденциальности. Системы менеджмента информационной безопасности. Требования. – М.: Стандартинформ, 2022. – 28 с.

28. ГОСТ Р 59276-2020. Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения. – М.: Стандартинформ, 2021. – 16 с.

29. Page M.J., McKenzie J.E., Bossuyt P.M. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews // BMJ. – 2021. – Vol. 372. – Article n71. – DOI: 10.1136/bmj.n71.

30. Bai Y., Mei J., Yuille A.L., Xie C. Are transformers more robust than CNNs? // Advances in Neural Information Processing Systems (NeurIPS). – 2021. – Vol. 34. – arXiv:2111.05464 [Электронный ресурс]. – URL: https://arxiv.org/abs/2111.05464 (дата обращения: 15.05.2026).

31. Gowal S., Rebuffi S.-A., Wiles O., Stimberg F., Calian D.A., Mann T. Improving robustness using generated data // Advances in Neural Information Processing Systems (NeurIPS). – 2021. – Vol. 34. – arXiv:2110.09468 [Электронный ресурс]. – URL: https://arxiv.org/abs/2110.09468 (дата обращения: 15.05.2026).

32. Rebuffi S.-A., Gowal S., Calian D.A., Stimberg F., Wiles O., Mann T. Fixing data augmentation to improve adversarial robustness // arXiv preprint. – 2021. – arXiv:2103.01946 [Электронный ресурс]. – URL: https://arxiv.org/abs/2103.01946 (дата обращения: 15.05.2026).

33. Sehwag V., Mahloujifar S., Handina T., Dai S., Xiang C., Chiang M., Mittal P. Robust learning meets generative models: Can proxy distributions improve adversarial robustness? // International Conference on Learning Representations (ICLR). – 2022. – arXiv:2104.09425 [Электронный ресурс]. – URL: https://arxiv.org/abs/2104.09425 (дата обращения: 15.05.2026).

34. ISO/IEC 23894:2023. Information technology – Artificial intelligence – Guidance on risk management. – Geneva: International Organization for Standardization, 2023. — 30 p.

35. NIST AI Risk Management Framework (AI RMF 1.0). – Gaithersburg: National Institute of Standards and Technology, 2023. – 48 p. – DOI: 10.6028/NIST.AI.100-1.

36. Банк данных угроз безопасности информации [Электронный ресурс] // ФСТЭК России. – URL: https://bdu.fstec.ru/ (дата обращения: 15.05.2026).

37 Намиот Д.Е., Ильюшин Е.А. Атаки на системы машинного обучения // International Journal of Open Information Technologies. – 2022. – Т. 10, № 3. – С. 57–66.

References

1. LeCun Y., Bengio Y., Hinton G. Deep learning. Nature, 2015, vol. 521, no. 7553, pp. 436–444. DOI: 10.1038/nature14539.

2. Szegedy C., Zaremba W., Sutskever I., Bruna J., Erhan D., Goodfellow I., Fergus R. Intriguing properties of neural networks. Proc. ICLR, 2014, arXiv:1312.6199. Available at: https://arxiv.org/abs/1312.6199 (accessed 15.05.2026).

3. Goodfellow I.J., Shlens J., Szegedy C. Explaining and harnessing adversarial examples. Proc. ICLR, 2015, arXiv:1412.6572. Available at: https://arxiv.org/abs/1412.6572 (accessed 15.05.2026).

4. Biggio B., Roli F. Wild patterns: Ten years after the rise of adversarial machine learning. Pattern Recognition, 2018, vol. 84, pp. 317–331. DOI: 10.1016/j.patcog.2018.07.023.

5. Akhtar N., Mian A. Threat of adversarial attacks on deep learning in computer vision: A survey. IEEE Access, 2018, vol. 6, pp. 14410–14430. DOI: 10.1109/ACCESS.2018.2807385.

6. Moosavi-Dezfooli S.-M., Fawzi A., Fawzi O., Frossard P. Universal adversarial perturbations. Proc. IEEE CVPR, 2017, pp. 1765–1773. DOI: 10.1109/CVPR.2017.17.

7. Carlini N., Wagner D. Towards evaluating the robustness of neural networks. Proc. IEEE S&P, 2017, pp. 39–57. DOI: 10.1109/SP.2017.49.

8. Finlayson S.G., Bowers J.D., Ito J., Zittrain J.L., Beam A.L., Kohane I.S. Adversarial attacks on medical machine learning. Science, 2019, vol. 363, no. 6433, pp. 1287–1289. DOI: 10.1126/science.aaw4399.

9. Tramèr F., Papernot N., Goodfellow I., Boneh D., McDaniel P. The space of transferable adversarial examples. arXiv preprint, 2017, arXiv:1704.03453. Available at: https://arxiv.org/abs/1704.03453 (accessed 15.05.2026).

10. Eykholt K., Evtimov I., Fernandes E., Li B., Rahmati A., Xiao C., Prakash A., Kohno T., Song D. Robust physical-world attacks on deep learning visual classification. Proc. IEEE CVPR, 2018, pp. 1625–1634. DOI: 10.1109/CVPR.2018.00175.

11. Mopuri K.R., Ojha U., Garg U., Babu R.V. NAG: Network for adversary generation. Proc. IEEE CVPR, 2018, pp. 742–751. DOI: 10.1109/CVPR.2018.00084.

12. Madry A., Makelov A., Schmidt L., Tsipras D., Vladu A. Towards deep learning models resistant to adversarial attacks. Proc. ICLR, 2018, arXiv:1706.06083. Available at: https://arxiv.org/abs/1706.06083 (accessed 15.05.2026).

13. Croce F., Andriushchenko M., Sehwag V., Debenedetti E., Flammarion N., Chiang M., Mittal P., Hein M. RobustBench: A standardized adversarial robustness benchmark. Proc. NeurIPS Datasets and Benchmarks Track, 2021, arXiv:2010.09670. Available at: https://arxiv.org/abs/2010.09670 (accessed 15.05.2026).

14. Athalye A., Carlini N., Wagner D. Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples. Proc. ICML, 2018, pp. 274–283, arXiv:1802.00420. Available at: https://arxiv.org/abs/1802.00420 (accessed 15.05.2026).

15. Papernot N., McDaniel P., Wu X., Jha S., Swami A. Distillation as a defense to adversarial perturbations against deep neural networks. Proc. IEEE S&P, 2016, pp. 582–597. DOI: 10.1109/SP.2016.41.

16. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial nets. Proc. NeurIPS, 2014, vol. 27, arXiv:1406.2661. Available at: https://arxiv.org/abs/1406.2661 (accessed 15.05.2026).

17. Wong E., Rice L., Kolter J.Z. Fast is better than free: Revisiting adversarial training. Proc. ICLR, 2020, arXiv:2001.03994. Available at: https://arxiv.org/abs/2001.03994 (accessed 15.05.2026).

18. Zhang H., Yu Y., Jiao J., Xing E.P., El Ghaoui L., Jordan M.I. Theoretically principled trade-off between robustness and accuracy. Proc. ICML, 2019, pp. 7472–7482, arXiv:1901.08573. Available at: https://arxiv.org/abs/1901.08573 (accessed 15.05.2026).

19. Shafahi A., Najibi M., Ghiasi M.A., Xu Z., Dickerson J., Studer C., Davis L.S., Taylor G., Goldstein T. Adversarial training for free! Proc. NeurIPS, 2019, vol. 32, arXiv:1904.12843. Available at: https://arxiv.org/abs/1904.12843 (accessed 15.05.2026).

20. Cohen J., Rosenfeld E., Kolter J.Z. Certified adversarial robustness via randomized smoothing. Proc. ICML, 2019, pp. 1310–1320, arXiv:1902.02918. Available at: https://arxiv.org/abs/1902.02918 (accessed 15.05.2026).

21. Croce F., Hein M. Reliable evaluation of adversarial robustness with an ensemble of diverse parameter-free attacks. Proc. ICML, 2020, pp. 2206–2216, arXiv:2003.01690. Available at: https://arxiv.org/abs/2003.01690 (accessed 15.05.2026).

22. Hendrycks D., Dietterich T. Benchmarking neural network robustness to common corruptions and perturbations. Proc. ICLR, 2019, arXiv:1903.12261. Available at: https://arxiv.org/abs/1903.12261 (accessed 15.05.2026).

23. Prikaz FSTEK Rossii ot 11.02.2013 No. 17 «Ob utverzhdenii Trebovaniy o zashchite informatsii, ne sostavlyayushchey gosudarstvennuyu taynu, soderzhashcheysya v gosudarstvennykh informatsionnykh sistemakh» (s izm.) [Order of the FSTEC of Russia No. 17 of 11.02.2013 "On approval of the Requirements for the protection of information not constituting a state secret contained in state information systems" (as amended)]. FSTEC of Russia. Available at: https://fstec.ru/ (accessed 15.05.2026). (In Russian).

24. Prikaz FSTEK Rossii ot 18.02.2013 No. 21 «Ob utverzhdenii Sostava i soderzhaniya organizatsionnykh i tekhnicheskikh mer po obespecheniyu bezopasnosti personal'nykh dannykh pri ikh obrabotke v informatsionnykh sistemakh personal'nykh dannykh» (s izm.) [Order of the FSTEC of Russia No. 21 of 18.02.2013 "On approval of the Composition and content of organizational and technical measures to ensure the security of personal data during its processing in personal data information systems" (as amended)]. FSTEC of Russia. Available at: https://fstec.ru/ (accessed 15.05.2026). (In Russian).

25. Prikaz FSTEK Rossii ot 25.12.2017 No. 239 «Ob utverzhdenii Trebovaniy po obespecheniyu bezopasnosti znachimykh ob"yektov kriticheskoy informatsionnoy infrastruktury Rossiyskoy Federatsii» (s izm.) [Order of the FSTEC of Russia No. 239 of 25.12.2017 "On approval of the Requirements for ensuring the security of significant objects of the critical information infrastructure of the Russian Federation" (as amended)]. FSTEC of Russia. Available at: https://fstec.ru/ (accessed 15.05.2026). (In Russian).

26. Federal'nyy zakon ot 26.07.2017 No. 187-FZ «O bezopasnosti kriticheskoy informatsionnoy

infrastruktury Rossiyskoy Federatsii» [Federal Law No. 187-FZ of 26.07.2017 "On the security of the critical information infrastructure of the Russian Federation"]. Sobranie zakonodatel'stva Rossiyskoy Federatsii, 2017, no. 31 (pt. I), art. 4736. (In Russian).

27. GOST R ISO/MEK 27001-2021. Informatsionnaya bezopasnost', kiberbezopasnost' i zashchita konfidentsial'nosti. Sistemy menedzhmenta informatsionnoy bezopasnosti. Trebovaniya [GOST R ISO/IEC 27001-2021. Information security, cybersecurity and privacy protection. Information security management systems. Requirements]. Moscow, Standartinform Publ., 2022. 28 p. (In Russian).

28. GOST R 59276-2020. Sistemy iskusstvennogo intellekta. Sposoby obespecheniya doveriya. Obshchie polozheniya [GOST R 59276-2020. Artificial intelligence systems. Methods for ensuring trust. General provisions]. Moscow, Standartinform Publ., 2021. 16 p. (In Russian).

29. Page M.J., McKenzie J.E., Bossuyt P.M. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, 2021, vol. 372, n71. DOI: 10.1136/bmj.n71.

30. Bai Y., Mei J., Yuille A.L., Xie C. Are transformers more robust than CNNs? Proc. NeurIPS, 2021, vol. 34, arXiv:2111.05464. Available at: https://arxiv.org/abs/2111.05464 (accessed 15.05.2026).

31. Gowal S., Rebuffi S.-A., Wiles O., Stimberg F., Calian D.A., Mann T. Improving robustness using generated data. Proc. NeurIPS, 2021, vol. 34, arXiv:2110.09468. Available at: https://arxiv.org/abs/2110.09468 (accessed 15.05.2026).

32. Rebuffi S.-A., Gowal S., Calian D.A., Stimberg F., Wiles O., Mann T. Fixing data augmentation to improve adversarial robustness. arXiv preprint, 2021, arXiv:2103.01946. Available at: https://arxiv.org/abs/2103.01946 (accessed 15.05.2026).

33. Sehwag V., Mahloujifar S., Handina T., Dai S., Xiang C., Chiang M., Mittal P. Robust learning meets generative models: Can proxy distributions improve adversarial robustness? Proc. ICLR, 2022, arXiv:2104.09425. Available at: https://arxiv.org/abs/2104.09425 (accessed 15.05.2026).

34. ISO/IEC 23894:2023. Information technology – Artificial intelligence – Guidance on risk management. Geneva, ISO, 2023. 30 p.

35. NIST AI Risk Management Framework (AI RMF 1.0). Gaithersburg, NIST, 2023. 48 p. DOI: 10.6028/NIST.AI.100-1.

36. Bank dannykh ugroz bezopasnosti informatsii [Information security threat data bank]. FSTEC of Russia. Available at: https://bdu.fstec.ru/ (accessed 15.05.2026). (In Russian).

37. Namiot D.E., Ilyushin E.A. Ataki na sistemy mashinnogo obucheniya [Attacks on machine learning systems]. International Journal of Open Information Technologies, 2022, vol. 10, no. 3, pp. 57–66. (In Russian).

ЧАСТИКОВА Вера Аркадьевна, кандидат технических наук, доцент, доцент кафедры кибербезопасности и защиты информации федерального государственного бюджетного образовательного учреждения высшего образования «Кубанский государственный технологический университет» (ФГБОУ ВО «КубГТУ»). 350072, г. Краснодар, ул. Московская, д. 2. E-mail: chastikova_va@mail.ru

МАРЧЕНКО Вячеслав Андреевич, магистрант кафедры кибербезопасности и защиты информации федерального государственного бюджетного образовательного учреждения высшего образования «Кубанский государственный технологический университет» (ФГБОУ ВО «КубГТУ»). 350072, г. Краснодар, ул. Московская, д. 2. E-mail: marche11a.slava@mail.ru.

CHASTIKOVA Vera Arkadievna, Candidate of Technical Sciences, Associate Professor of the Department of Cybersecurity and Information Protection at the Federal State Budgetary Educational Institution of Higher Edu-cation "Kuban State Technological University". 350072, Krasnodar, Moskovskaya str., 2. E-mail: chastikova_va@mail.ru

MARCHENKO Vyacheslav Andreevich, master's degree student of the Department of Cybersecurity and Information Protection at the Federal State Budgetary Educational Institution of Higher Edu-cation "Kuban State Technological University". 350072, Krasnodar, Moskovskaya str., 2. E-mail: marche11a.slava@mail.ru.

УДК 004.056.53 Вестник УрФО № 2(60) / 2026, с. 88–97

Ощепков Н.В.

DOI: 10.14529/secur260209

ОЦЕНКА ВЛИЯНИЯ МЕТОДОВ ВЕРИФИКАЦИИ ОТПРАВИТЕЛЯ НА БЫСТРОДЕЙСТВИЕ СИСТЕМЫ УПРАВЛЕНИЯ РЕЛЕЙНО-ПУСКОВЫМ БЛОКОМ ШТАНГОВОГО ГЛУБИННОГО НАСОСА ПРИ РАЗЛИЧНЫХ ТАКТОВЫХ ЧАСТОТАХ МИКРОКОНТРОЛЛЕРА

В статье выполнена оценка влияния методов верификации отправителя команд на быстродействие системы управления релейно-пусковым блоком штангового глубинного насоса, функционирующей на основе протокола Modbus RTU. Исследование проводилось непосредственно на объекте внедрения – скважине №309 куста №4 Чернушинского месторождения – при пяти тактовых частотах микроконтроллера ESP32, используемого в составе ведомого устройства: 20, 40, 80, 160 и 240 МГц, и трёх значениях битрейта: 9600, 19200 и 115200 бит/с, при длине линии RS-485 50 м. Показано, что увеличение битрейта и тактовой частоты микроконтроллера сужает разброс задержек доставки кадра, что позволяет обоснованно уменьшить длительность временного окна верификации и тем самым снизить вероятность успешной реализации атаки подмены мастера. Быстродействие системы в рабочем режиме определяется канальной задержкой RS-485, что обосновывает применимость разработанных методов на микроконтроллерных платформах с различными тактовыми характеристиками без ущерба для быстродействия системы управления.

Ключевые слова: *Modbus RTU, RS-485, верификация отправителя, XOR-аутентификация, HMAC-SHA1.*

EVALUATION OF THE IMPACT OF SENDER VERIFICATION METHODS ON THE PERFORMANCE OF THE CONTROL SYSTEM FOR A PUMP RELAY-START UNIT AT VARIOUS MICROCONTROLLER CLOCK FREQUENCIES

This article evaluates the impact of command-sender verification methods on the response speed of a control system for a pump relay-start unit operating via the Modbus RTU protocol. The study was conducted at the actual deployment site-Well No. 309, Cluster No. 4 of the Chernushinskoye oil field-using an ESP32 microcontroller as the slave device. Testing involved five clock frequencies (20, 40, 80, 160, and 240 MHz) and three bitrates (9,600, 19,200, and 115,200 bps) over a 50-meter RS-485 line. The results demonstrate that increasing the bitrate and microcontroller clock frequency reduces the variance in frame delivery latency; this allows for a justified reduction in the verification time window, thereby lowering the probability of a successful master-spoofing attack. Under normal operating conditions, system speed is determined by RS-485 channel latency, confirming the suitability of the developed methods for microcontroller platforms with varying clock speeds without compromising control system performance.

Keywords: *Modbus RTU, RS-485, sender verification, XOR authentication, HMAC-SHA1.*

Введение

Протокол Modbus RTU, разработанный компанией Modicon в 1979 году, по сей день остаётся одним из наиболее распространённых транспортных протоколов в промышленных системах управления – АСУ ТП, SCADA и «интернет вещей» (IoT) [1]. В нефтяной промышленности он традиционно используется для дистанционного управления приводным оборудованием скважин, в том числе релейно-пусковыми блоками штанговых глубинных насосов (ШГН). Вместе с тем протокол проектировался без учёта требований информационной безопасности: он не предусматривает ни механизмов аутентификации участников обмена, ни шифрования передаваемых данных.

Отсутствие аутентификации и шифрования в протоколе Modbus/TCP позволяет злоумышленнику реализовать атаку типа «человек посередине» (MITM), перехватывая и модифицируя трафик между SCADA-системой и программируемым логическим контроллером (ПЛК), а также внедряя ложные команды от имени легитимного мастера [2]. Практическая опасность таких уязвимостей подтверждается реальными инцидентами: в январе 2024 года вредоносная программа FrostyGoop, эксплуатирующая незащищённый протокол Modbus, была применена в кибератаке на предприятие теплоснабжения – более 600 жилых домов остались без отопления на двое суток при отрицательных температурах [3].

Katulić и соавторы предложили использо-

вание кодов аутентификации сообщений (MAC) для Modbus/TCP и экспериментально подтвердили их практическую применимость в промышленных сетях [4]. Для ресурсоограниченных устройств класса IIoT предложены лёгкие протоколы аутентификации на основе операций XOR и хэш-функций, обеспечивающие небольшое время выполнения при формально подтверждённой защищённости [5].

Калинкин Я.В. и Тутов И.А. разработали программно-аппаратный модуль фильтрации кадров Modbus RTU по белым спискам допустимых значений полей, однако данный подход не предусматривает верификации источника команды и не защищает от атаки подмены мастера [6].

Безукладников И.И. и Кон Е.Л. предложили универсальный подход к организации скрытых каналов в распределённых АСУ, проанализировали шлюз LonWorks–Internet и разработали способ организации скрытого канала в системе управления интеллектуальным зданием [7]. Капгер И.В., Сизов В.П. и Южаков А.А. разработали алгоритм потокового шифрования для промышленных сетей LON на основе функции косинуса с 48-битным ключом аутентификации и динамической сменой сессионных ключей каждые 256 циклов [8]. Капгер И.В. в диссертационном исследовании рассматривал проблему повышения уровня конфиденциальности в промышленных сетях систем автоматизации испытаний, разработав методы и средства криптографической защиты передаваемых данных применительно к ресурсоограниченным узлам промышленных сетей [9].

В работе Попова П.А., Розенберга Е.Н., Сабанова А.Г. и Шубинского И.Б. сформулированы концепции безопасности верхнего уровня управления АСУ ТП ЖТ как объектов критической информационной инфраструктуры (КИИ), включая концепции защищённой среды, взаимозависимости функциональной и информационной безопасности и единой модели угроз [10]. Dontha J.D. исследовала применение граничных вычислений в ICS, предложив многоуровневую архитектуру, объединяющую машинное обучение, блокчейн и шифрование для снижения задержек и повышения устойчивости к кибератакам [11].

Ádámkó É., Jakabóczki G. и Szemes P.T. предложили протокол защиты Modbus RTU на основе схемы разделения секрета Шамира с взаимной аутентификацией и AES-шифрованием без изменения структуры стандартного кадра, подтверждённый апробацией на реальном устройстве сбора данных [12]. Грешонков А.М. и Моисеев В.С. обосновали применение машинного обучения для выявления аномалий в IIoT-системах и предложили многоуровневую архитектуру системы безопасности с поддержкой протоколов Modbus, MQTT и OPC-UA [13].

Шигапов И.И., Ильмушкина А.А. и Юсупова А.С. разработали концептуальную архитектуру самоорганизующихся IIoT-сетей на базе IEEE 802.15.4/802.15.5, интегрирующую эллиптическую криптографию, пороговые схемы Шамира и модель управления доверием на основе риск-скора [14]. Patel H.S., Gautam C.S. и Waoo A.A. экспериментально показали на выборке из 120 устройств, что blockchain-механизмы и машинное обучение обеспечивают наивысшую эффективность обнаружения атак – 96% и 94% соответственно [15]. Santhiya R. и Barakath Begam A. установили, что интеграция SDN и провайдеров управления безопасностью существенно повышает защищённость IoT-устройств «умного дома», тогда как протоколы MQTT и CoAP требуют обязательного дополнения уровнем TLS/DTLS [16]. Таржанов Т.В. и Новиков С.Н. выполнили имитационное моделирование DoS-атаки на виртуальную IoT-сеть и подтвердили эффективность защиты посредством межсетевого экранирования средствами POX-контроллера [17].

Таким образом, анализ литературы показывает, что тематика защиты промышленных коммуникационных протоколов – в первую очередь Modbus RTU – в контексте АСУ ТП, IoT и IIoT вызывает устойчивый исследовательский интерес: работы охватывают широкий спектр подходов от лёгких криптографических протоколов и схем разделения секрета до методов машинного обучения и архитектур граничных вычислений, однако специфика последовательной шины RS-485 с жёсткими ресурсными ограничениями ведомых устройств и проблема верификации отправителя команд остаются недостаточно изученными областями.

Ранее автором был разработан и внедрён механизм верификации отправителя команд для системы управления релейно-пусковым блоком ШГН на скважине №309 куста №4 Чернушинского месторождения предприятия ООО «РОССМА» [18]. Механизм включает три компонента: метод XOR-аутентификации с 16-битным ключом [19], метод верификации на основе периодически открываемых вре-

менных окон с криптографическим согласованием параметров на базе HMAC-SHA1 [20], а также их комбинацию. По результатам натурных испытаний вероятность успешной реализации атаки подмены мастера при комбинированном методе составила $P \approx 3{,}27 \cdot 10^{-6}$.

В ходе эксплуатации системы был поставлен практический вопрос: как ведут себя разработанные методы верификации при изменении тактовой частоты микроконтроллера ведомого устройства? ESP32, применяемый в составе релейно-пускового блока, допускает программное переключение тактовой частоты ядра между штатными значениями 80, 160 и 240 МГц, а также может эксплуатироваться на пониженных частотах (20 и 40 МГц) с целью снижения энергопотребления или тепловыделения. Расширение парка устройств на аналогичных объектах предполагает использование платформ с различными тактовыми характеристиками, что ставит задачу обоснования переносимости разработанных методов без перенастройки параметров верификации.

Представленные результаты настоящего исследования состоят в следующем:

1. Выполнена количественная оценка влияния тактовой частоты микроконтроллера ESP32 (20 – 240 МГц) и битрейта RS-485 (9600 – 115 200 бит/с) на джиттер задержки доставки кадра Modbus RTU в условиях реальной промышленной эксплуатации.

2. Обоснована возможность динамического сужения параметра D (длительности временного окна верификации) на основе измеренного диапазона джиттера канала без потери надёжности приёма легитимных кадров.

3. Показана применимость разработанных методов верификации на микроконтроллерных платформах с различными тактовыми характеристиками без ущерба для быстродействия системы управления.

Объектом исследования является система дистанционного управления приводным комплексом (СДУПК) ШГН на скважине №309 куста №4 Чернушинского месторождения, подробно описанная в [18]. Система построена по топологии шины с протоколом «ведущий-ведомый». Ведущим устройством является АРМ оператора на базе ПК с адаптером USB-RS-485 [21]. Ведомыми устройствами выступают датчики и управляющее устройство релейно-пускового блока на базе микроконтроллера ESP32 под управлением MicroPython. Управление состоянием привода насоса осуществляется командой записи дискретного выхода с адресом 0x0000 посредством функционального кода FC05. Физическая среда передачи – двухпроводная линия RS-485, длина шины 50 м. Структурная схема СДУПК приведена на рис. 1. Система включает в себя ведущее устройство (мастер) и 6 ведомых устройств: датчик подсчёта количества наклонов, релейно-пусковой блок, динамограф, датчик давления, датчик температуры двигателя и датчик угла наклона.

Материалы и методы

Механизм верификации отправителя реализован в прошивке ESP32 ведомого устройства релейно-пускового блока и включает два метода, применяемых совместно.

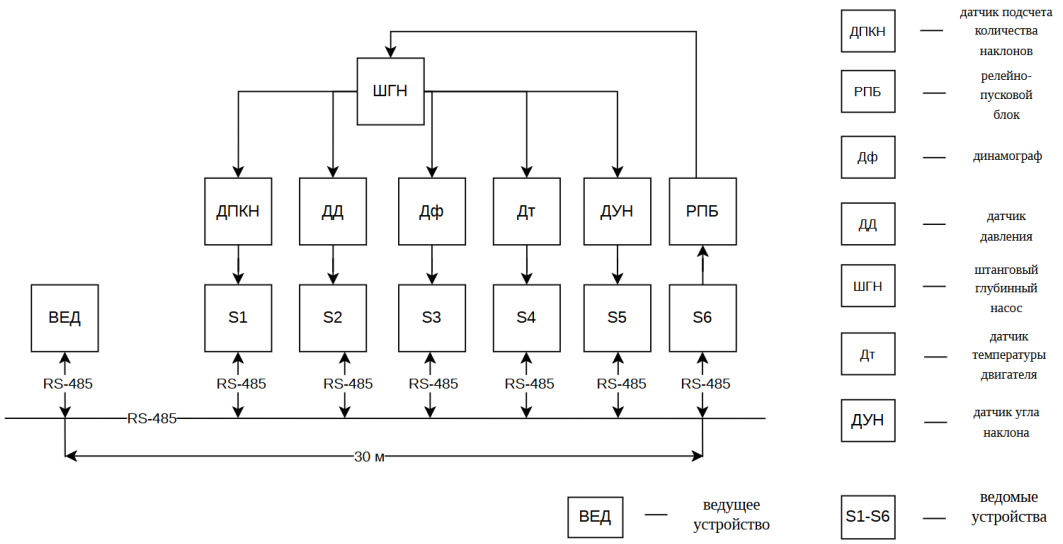

Рис. 1. Структурная схема системы дистанционного управления ШГН скважины

Метод XOR-аутентификации, при котором поле данных каждого исходящего кадра FC05 гаммируется мастером с использованием 16-битного разделяемого ключа K посредством операции побитового исключающего ИЛИ. Ведомое устройство при приёме выполняет обратную операцию и проверяет корректность результата. Вероятность того, что случайный кадр злоумышленника пройдёт XOR-проверку, равна $3{,}05 \cdot 10^{-5}$ (при заданном ключе равном 16 бит).

Метод временных окон. Ведомое устройство принимает кадры только в течение периодически открываемых временных окон. Параметры окон передаются мастером в кадре контракта с использованием функциональных кодов из пользовательского диапазона спецификации Modbus (65 – 72) с криптографическим согласованием на базе HMAC-SHA1 [20]. Параметр D задаёт длину временного окна: легитимный кадр считается действительным, если он поступил в интервале [0; D] от момента открытия окна. Параметр T задаёт длительность закрытого интервала между окнами. Полный период составляет D + T. Вероятность того, что произвольный кадр злоумышленника попадёт в открытое окно:

$$P_W = \frac{D}{T+D}. \quad (1)$$

Комбинированный метод является сочетанием обоих методов. События «кадр попал в окно» и «кадр прошёл XOR-проверку» независимы, поэтому вероятность успешной атаки подмены мастера:

$$P_{W\_XOR} = \frac{D}{2^{15} \cdot (T+D)}. \quad (2)$$

При подстановке исходных параметров объекта внедрения [18] D = 30 мс, T = 250 мс в формулу (2) вероятность реализации атаки подмены и отправки кадра на ведомое устройство составляет $3{,}27 \cdot 10^{-6}$.

Методика измерений

Для оценки влияния тактовой частоты микроконтроллера и битрейта на джиттер доставки кадра измерения проводились непосредственно на объекте внедрения при трёх значениях битрейта (9600, 19200 и 115200 бит/с) и пяти тактовых частотах ESP32 (20, 40, 80, 160 и 240 МГц). Переключение тактовой частоты осуществлялось программно средствами MicroPython (machine.freq()). Для каждой из 15 комбинаций параметров формировалась выборка объёмом n = 1000 измерений задержки доставки кадра в установившемся режиме работы системы.

По каждой выборке фиксировался наблюдаемый диапазон задержек [D_{min}, D_{max}]. Фиксация значения момента времени попадания кадра мастера на ведомое устройство в рамках временного интервала приема сообщения производилась ведомым устройством благодаря встроенной в MicroPython функции time.ticks_us в момент получения первого байта кадра. На основе этого диапазона вычислялось скорректированное минимально допустимое значение параметра $D_{корр}$. Скорректированное значение $D_{корр}$ вычислялось как разность между исходным параметром окна D = 30 мс и наблюдаемым минимумом задержки D_{min} с добавлением страхового запаса в 2 – 3 мс для исключения ложных отказов при приёме легитимных кадров.

Результаты

Анализ данных, представленных в табл.1 и на рис. 2 – 4, позволяет выявить ряд закономерностей. Наиболее существенное влияние на величину джиттера оказывает увеличение битрейта: с ростом скорости передачи диапазон его значений уменьшается. Это связано с тем, что при более высоком битрейте время передачи каждого байта сокращается, вследствие чего уменьшается общая длительность передачи кадра и снижается временной разброс его доставки. Так, при скорости передачи 19200 бит/с становится возможным уменьшение длительности окна D с 30 до 12 мс. В результате при значении T = 250 мс вероятность успешной реализации атаки подмены кадра, рассчитанная по формуле (2), снижается до $1{,}40 \cdot 10^{-6}$.

Во-вторых, влияние тактовой частоты микроконтроллера на диапазон джиттера проявляется по-разному в зависимости от битрейта. При 19200 бит/с – рабочей скорости объекта внедрения – диапазон [D_{min}, D_{max}] практически не изменяется при смене частоты от 20 до 240 МГц, а значение $D_{корр}$ = 12 мс одинаково для всех пяти частот. При 9600 бит/с повышение частоты с 20 до 240 МГц позволяет сократить $D_{корр}$ с 25 до 22 мс. Наиболее выраженный эффект наблюдается при 115200 бит/с: повышение частоты с 20 до 240 МГц сокращает $D_{корр}$ с 7 до 4 мс – сужение почти вдвое.

Обсуждение

Наблюдаемые закономерности объясняются соотношением двух составляющих суммарной задержки доставки кадра: канальной задержки, определяемой временем побитовой передачи по линии RS-485, и программ-

Таблица 1

Диапазон задержек доставки кадра и скорректированное значение Dкорр при различных битрейтах и тактовых частотах ESP32

Битрейт, бит/с	Частота, МГц	D_{min} – D_{max}, мс	$D_{корр}$, мс
9600	20	6 – 20	25
	40	7 – 19	24
	80	8 – 18	23
	160	9 – 17	22
	240	9 – 16	22
19 200	20	18 – 25	12
	40	19 – 25	12
	80	19 – 25	12
	160	19 – 25	12
	240	20 – 25	12
115 200	20	25 – 29	7
	40	26 – 29	6
	80	27 – 29	5
	160	28 – 29	5
	240	28 – 29	4

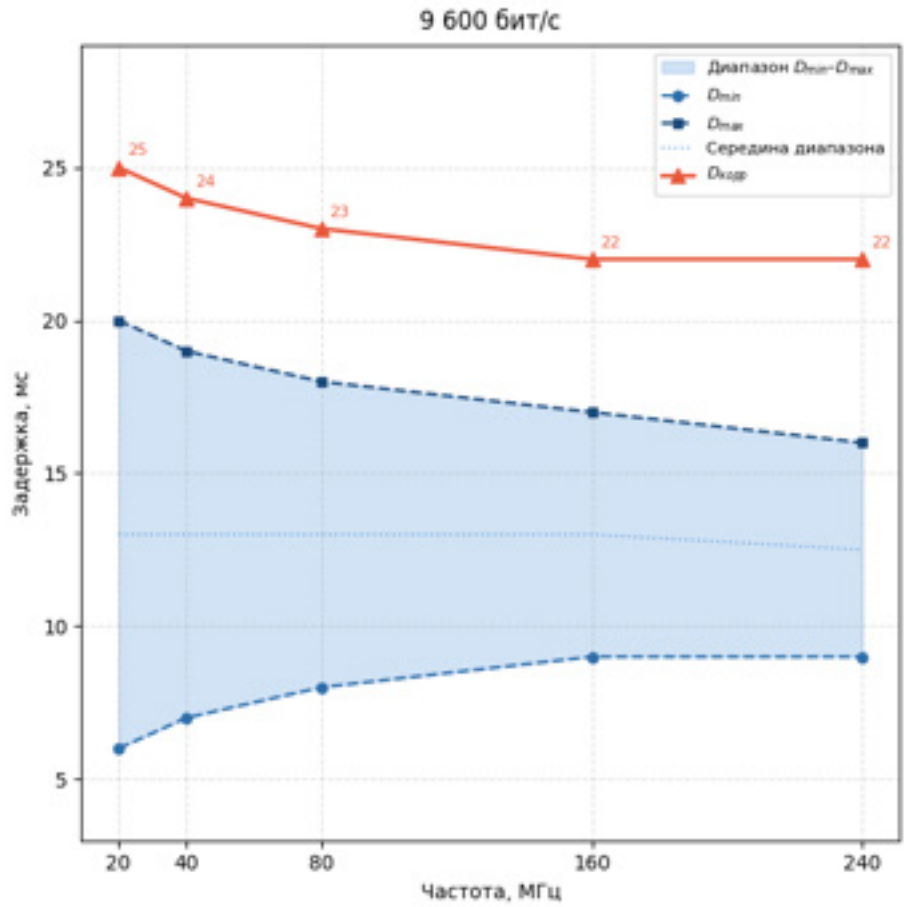

Рис. 2. Распределение задержек Modbus по битрейту при 9600 бит /с

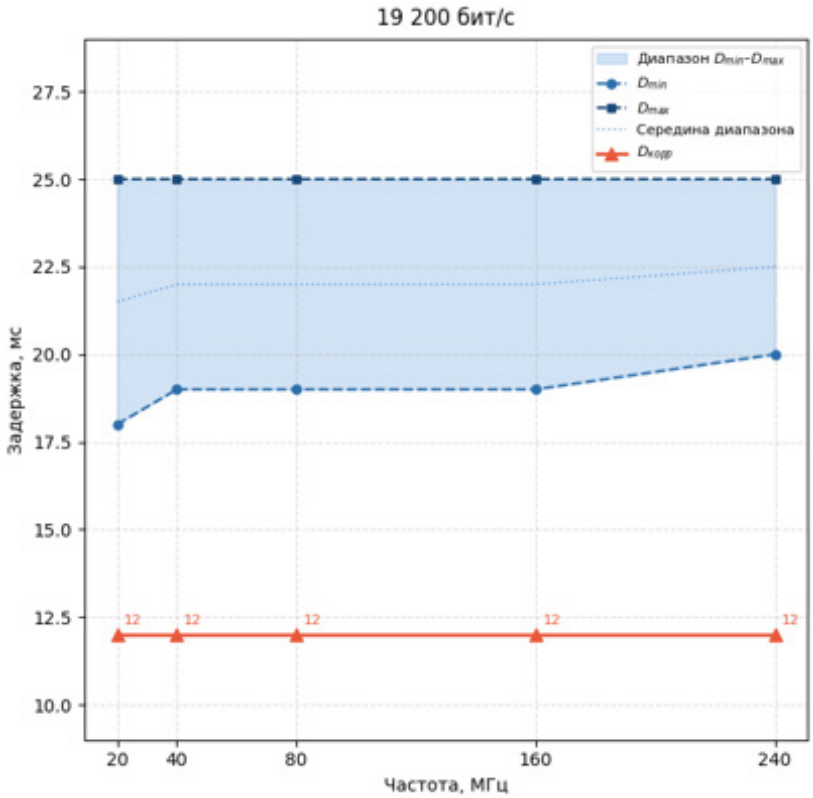

Рис. 3. Распределение задержек Modbus по битрейту при 19200 бит /с

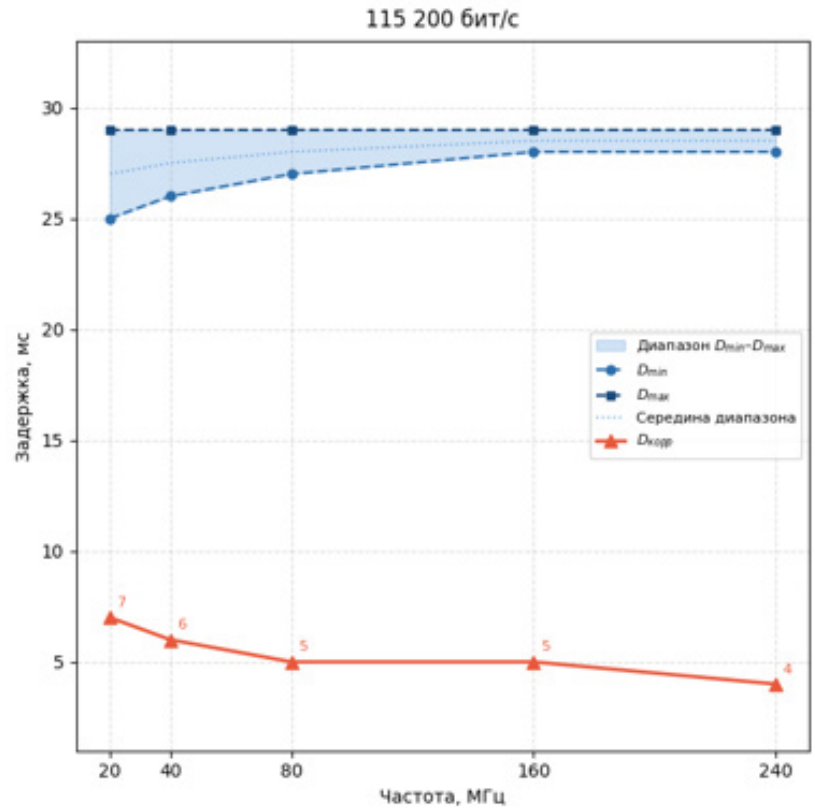

Рис. 4. Распределение задержек Modbus по битрейту при 115200 бит /с

ной задержки, обусловленной временем реакции стека MicroPython на прерывание UART.

При битрейте 19200 бит/с время передачи кадра по каналу является доминирующей составляющей задержки и полностью маскирует вариацию программной задержки обработки прерывания. Независимо от тактовой частоты ядра ESP32 суммарный джиттер определяется физическим каналом RS-485, что и объясняет постоянство $D_{корр}$ = 12 мс во всём диапазоне 20 – 240 МГц.

При битрейте 115200 бит/с время передачи кадра сокращается примерно в шесть раз по сравнению с 19200 бит/с, и программная составляющая задержки становится сопоставимой с канальной. В этом режиме тактовая частота микроконтроллера начинает существенно влиять на джиттер, что и даёт наблюдаемый эффект двукратного сужения $D_{корр}$ при увеличении частоты с 20 до 240 МГц.

Заключение

При рабочей скорости объекта внедрения 19200 бит/с значение $D_{корр}$ = 12 мс не зависит от тактовой частоты микроконтроллера в диапазоне 20 – 240 МГц. Это означает, что исходный параметр окна D = 30 мс, установленный при внедрении, может быть безопасно сокращён до 12 мс без риска ложных отказов при приёме легитимных кадров и без перенастройки при смене тактовой частоты. Сокращение окна с 30 до 12 мс уменьшает вероятность успешной атаки подмены мастера с $3{,}27 \cdot 10^{-6}$ до $1{,}40 \cdot 10^{-6}$, то есть более чем вдвое, при этом не требуя ни повторных натурных измерений, ни изменений в алгоритме верификации, ни модификации аппаратной конфигурации системы. Дополнительным следствием является возможность перевода микроконтроллера на пониженную тактовую частоту – например, с 240 МГц до 80 МГц или ниже – с целью снижения энергопотребления или тепловыделения в условиях промышленной эксплуатации, без какого-либо ущерба для корректности работы механизма верификации при данном битрейте. Следует отметить, что при битрейте 115 200 бит/с тактовая частота микроконтроллера уже оказывает заметное влияние на джиттер, и повышение частоты с 20 до 240 МГц позволяет сократить $D_{корр}$ почти вдвое – с 7 до 4 мс. Это свидетельствует о том, что при наличии возможности дальнейшего увеличения тактовой частоты микроконтроллера сверх 240 МГц джиттер мог бы сузиться ещё больше, что позволило бы дополнительно уменьшить параметр D и тем самым повысить защищённость системы от атаки подмены мастера без каких-либо изменений в алгоритме верификации.

Литература

1. Modbus Protocol Specification [Электронный ресурс]. – Режим доступа: https://www.modbus.org/file/secure/modbusprotocolspecification.pdf. – Дата обращения: 05.06.2026.

2. Parian C., Guldimann T., Bhatia S. Fooling the Master: Exploiting Weaknesses in the Modbus Protocol [Электронный ресурс] // Procedia Computer Science. – 2020. – Vol. 171. – P. 2453–2458. – DOI: 10.1016/j.procs.2020.04.265. – URL: https://www.sciencedirect.com/science/article/pii/S1877050920312576 (дата обращения: 10.06.2026).

3. Нефёдова М. Вредонос FrostyGoop нацелен на промышленные системы управления [Электронный ресурс] // Хакер. – 2024. – 24 июля. – URL: https://xakep.ru/2024/07/24/frostygoop/ (дата обращения: 10.06.2026).

4. Katulić F., Sumina D., Groš S., Erceg I. Protecting Modbus/TCP-Based Industrial Automation and Control Systems Using Message Authentication Codes // IEEE Access. – 2023. – Vol. 11. – P. 47007–47023. – DOI: 10.1109/ACCESS.2023.3274592. – URL: https://ieeexplore.ieee.org/document/10122945 (дата обращения: 10.06.2026).

5. Lara E., Aguilar L., Sanchez M.A., García J.A. Lightweight Authentication Protocol for M2M Communications of Resource-Constrained Devices [Электронный ресурс] // Sensors. – 2020. – Vol. 20, No. 2. – Art. no. 501. – DOI: 10.3390/s20020501. – URL: https://www.mdpi.com/1424-8220/20/2/501 (дата обращения: 10.06.2026).

6. Калинкин Я.В., Тутов И.А. Модуль информационной безопасности в сетях Modbus RTU // Электронные средства и системы управления: материалы XVIII Международной научно-практической конференции, 16–18 ноября 2022 г. – Томск, 2022. – С. 57–60.

7. Безукладников И.И., Кон Е.Л. Скрытые каналы в распределенных автоматизированных системах // Вестник УГАТУ. – 2010. – Т. 14, № 2(37). – С. 245–250.

8. Капгер И.В., Сизов В.П., Южаков А.А. Криптографическое преобразование сообщений, передаваемых в промышленных сетях LON // Вопросы защиты информации. – 2010. – № 1(88). – С. 28–43.

9. Капгер И.В. Повышение уровня конфиденциальности в промышленных сетях систем автоматизации испытаний: дис. ... канд. техн. наук: 05.13.19. – Пермь, 2012. – 175 с.

10. Попов П.А., Розенберг Е.Н., Сабанов А.Г., Шубинский И.Б. Концепции обеспечения комплексной безопасности АСУ ТП верхнего уровня управления для объектов КИИ железнодорожного транспорта // Надёжность. – 2025. – Т. 25, № 3. – С. 42–49.

11. Dontha J.D. Edge Computing and Security in Industrial Control Systems // International Scientific Journal of Engineering and Management. – 2023. – Vol. 2, No. 9. – DOI: 10.55041/ISJEM01305.

12. Ádámkó É., Jakabóczki G., Szemes P.T. Proposal of a Secure Modbus RTU Communication with Adi Shamir›s Secret Sharing Method // International Journal of Electronics and Telecommunications. – 2018. – Vol. 64, No. 2. – P. 107–114. – DOI: 10.24425/119357.

13. Грешонков А.М., Моисеев В.С. Интеллектуальные системы обеспечения информационной безопасности в промышленном интернете вещей (IIoT) // Регион: системы, экономика, управление. – 2026. – № 4(71). – С. 210–211.

14. Шигапов И.И., Ильмушкина А.А., Юсупова А.С. Самоорганизующиеся беспроводные сети IIoT с учётом приоритетов безопасности // Экономика и управление: проблемы, решения. – 2025. – Т. 8, № 6(159). – С. 71–78.

15. Patel H.S., Gautam C.S., Waoo A.A. Advanced Security Vulnerabilities and Defense Mechanisms in IoT Networks // International Scientific Journal of Engineering and Management. – 2026. – Vol. 5, No. 6.

16. Santhiya R., Barakath Begam A. IoT Network Security in Smart Homes // International Scientific Journal of Engineering and Management. – 2026. – Vol. 5, No. 4.

17. Таржанов Т.В., Новиков С.Н. Исследование угроз информационной безопасности в приложениях IoT и методов защиты от этих угроз // Интерэкспо Гео-Сибирь. – 2021. – Т. 7, № 2. – С. 184–189.

18. Ощепков Н.В. механизм верификации отправителя команд в протоколе Modbus RTU: разработка и апробация на приводном оборудовании штангового глубинного насоса // Научно-технический вестник Поволжья. Научный журнал. №4, 2026. – С. 253-259.

19. Исключающее «или» // Википедия: [сайт]. – URL: https://ru.wikipedia.org/wiki/Исключающее_«или» (дата обращения: 09.06.2026).

20. RFC 4226: HOTP: An HMAC-Based One-Time Password Algorithm [Электронный ресурс]. – Режим доступа: https://www.rfc-editor.org/info/rfc4226. – Дата обращения: 05.06.2026.

21. RS-485 // Википедия: [сайт]. – URL: https://ru.wikipedia.org/wiki/RS-485 (дата обращения: 06.06.2026).

References

1. Modbus Protocol Specification [Electronic resource]. – Mode of access: https://www.modbus.org/file/secure/modbusprotocolspecification.pdf. – Date of access: 05.06.2026.

2. Parian C., Guldimann T., Bhatia S. Fooling the Master: Exploiting Weaknesses in the Modbus Protocol // Procedia Computer Science. – 2020. – Vol. 171. – P. 2453–2458. – DOI: 10.1016/j.procs.2020.04.265.

3. Nefyodova M. FrostyGoop malware targets industrial control systems [Vredonos FrostyGoop natseleen na promyshlennye sistemy upravleniya] // Haker [Hacker]. – 2024. – 24 July. – URL: https://xakep.ru/2024/07/24/frostygoop/ (date of access: 10.06.2026).

4. Katulić F., Sumina D., Groš S., Erceg I. Protecting Modbus/TCP-Based Industrial Automation and Control Systems Using Message Authentication Codes // IEEE Access. – 2023. – Vol. 11. – P. 47007–47023. – DOI: 10.1109/ACCESS.2023.3274592.

5. Lara E., Aguilar L., Sanchez M.A., García J.A. Lightweight Authentication Protocol for M2M Communications of Resource-Constrained Devices // Sensors. – 2020. – Vol. 20, No. 2. – Art. no. 501. – DOI: 10.3390/s20020501.

6. Kalinkin Ya.V., Tutov I.A. Information security module for Modbus RTU networks [Modul informatsionnoy bezopasnosti v setyakh Modbus RTU] // Elektronnye sredstva i sistemy upravleniya [Electronic Tools and Control Systems]: proceedings of the XVIII International Scientific and Practical Conference, 16–18 November 2022, Tomsk. – Tomsk, 2022. – P. 57–60.

7. Bezukladnikov I.I., Kon E.L. Covert channels in distributed automated systems [Skrytye kanaly v raspredelennykh avtomatizirovannykh sistemakh] // Vestnik Ufimskogo gosudarstvennogo aviatsionnogo tekhnicheskogo universiteta [Bulletin of Ufa State Aviation Technical University]. – 2010. – Vol. 14, No. 2(37). – P. 245–250.

8. Kapger I.V., Sizov V.P., Yuzhakov A.A. Cryptographic transformation of messages transmitted in LON industrial networks [Kriptograficheskoe preobrazovanie soobshcheniy, peredavaemykh v promyshlennykh setyakh LON] // Voprosy zashchity informatsii [Information Security Issues]. – 2010. – No. 1(88). – P. 28–43.

9. Kapger I.V. Improving the level of confidentiality in industrial networks of test automation systems: dissertation ... candidate of technical sciences: 05.13.19 [Povyshenie urovnya konfidentsialnosti v promyshlennykh setyakh sistem avtomatizatsii ispytaniy]. – Perm, 2012. – 175 p.

10. Popov P.A., Rozenberg E.N., Sabanov A.G., Shubinskiy I.B. Concepts of ensuring comprehensive security of upper-level control systems for critical information infrastructure objects of railway transport [Kontseptsii obespecheniya kompleksnoy bezopasnosti ASU TP verhnego urovnya upravleniya dlya obektov KII zheleznodorozhnogo transporta] // Nadyozhnost [Reliability]. – 2025. – Vol. 25, No. 3. – P. 42–49.

11. Dontha J.D. Edge Computing and Security in Industrial Control Systems // International Scientific Journal of Engineering and Management. – 2023. – Vol. 2, No. 9. – DOI: 10.55041/ISJEM01305.

12. Ádámkó É., Jakabóczki G., Szemes P.T. Proposal of a Secure Modbus RTU Communication with Adi Shamir›s Secret Sharing Method // International Journal of Electronics and Telecommunications. – 2018. – Vol. 64, No. 2. – P. 107–114. – DOI: 10.24425/119357.

13. Greshonkov A.M., Moiseev V.S. Intelligent systems for ensuring information security in the Industrial Internet of Things (IIoT) [Intellektualnye sistemy obespecheniya informatsionnoy bezopasnosti v promyshlennom interne veshchey (IIoT)] // Region: sistemy, ekonomika, upravlenie [Region: Systems, Economics, Management]. – 2026. – No. 4(71). – P. 210–211.

14. Shigapov I.I., Ilmushkina A.A., Yusupova A.S. Self-organizing wireless IIoT networks taking into account security priorities [Samorganizuyushchiesya besprovodnye seti IIoT s uchyotom prioritetov bezopasnosti] // Ekonomika i upravlenie: problemy, resheniya [Economics and Management: Problems, Solutions]. – 2025. – Vol. 8, No. 6(159). – P. 71–78.

15. Patel H.S., Gautam C.S., Waoo A.A. Advanced Security Vulnerabilities and Defense Mechanisms in IoT Networks // International Scientific Journal of Engineering and Management. – 2026. – Vol. 5, No. 6.

16. Santhiya R., Barakath Begam A. IoT Network Security in Smart Homes // International Scientific Journal of Engineering and Management. – 2026. – Vol. 5, No. 4.

17. Tarzhanov T.V., Novikov S.N. Study of information security threats in IoT applications and methods of protection against these threats [Issledovanie ugroz informatsionnoy bezopasnosti v prilozheniyakh IoT i metodov zashchity ot etikh ugroz] // InterEkspo Geo-Sibir. – 2021. – Vol. 7, No. 2. – P. 184–189.

18. Oshchepkov N.V. Sender verification mechanism for commands in the Modbus RTU protocol: development and testing on drive equipment of a sucker rod pump [Mekhanizm verifikatsii otpravitelya komand v protokole Modbus RTU: razrabotka i aprobatsiya na privodnom oborudovanii shtangovogo glubinnogo nasosa] // Nauchno-tekhnicheskiy vestnik Povolzhya [Scientific and Technical Bulletin of the Volga Region]. – 2026. – No. 4. – P. 253–259.

19. Exclusive or [Isklyuchayushcheye "ili"] // Wikipedia: the Free Encyclopedia [Vikipediya: Svobodnaya entsiklopediya]. – URL: https://ru.wikipedia.org/wiki/Исключающее_«или» (date of access: 09.06.2026).

20. RFC 4226: HOTP: An HMAC-Based One-Time Password Algorithm [Electronic resource]. – Mode of access: https://www.rfc-editor.org/info/rfc4226. – Date of access: 05.06.2026.

21. RS-485 // Wikipedia: the Free Encyclopedia [Vikipediya: Svobodnaya entsiklopediya]. – URL: https://ru.wikipedia.org/wiki/RS-485 (date of access: 06.06.2026).

ОЩЕПКОВ Никита Владимирович, аспирант кафедры автоматики и телемеханики, Пермский национальный исследовательский политехнический университет, ПНИПУ. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: maserati_2000@mail.ru

OSHCHEPKOV Nikita Vladimirovich, Postgraduate student of the Department of Automation and Telemechanics, Perm National Research Polytechnic University, 614990, Perm Region, Perm, Komsomolsky Prospekt, 29. E-mail: maserati_2000@mail.ru

Материалы к публикации отправлять по адресу E-mail: urvest@mail.ru
в редакцию журнала «Вестник УрФО. Безопасность в информационной сфере».

Или по почте по адресу: Россия, 454080, г. Челябинск, пр. им. Ленина, д. 76, ЮУрГУ,
Издательский центр

ВЕСТНИК УрФО
Безопасность в информационной сфере № 2(60) / 2026
Подписано в печать 30.06.2026. Дата выхода в свет 06.07.2026.
Формат 70×108 1/16. Печать цифровая. Усл.-печ. л. 5,78. Тираж 50 экз.
Заказ 128/245.
Цена свободная.

Отпечатано в типографии Издательского центра ФГАОУ ВО "ЮУрГУ (НИУ)".
454080, г. Челябинск, пр. им. В. И. Ленина, 76, ЮУрГУ, Издательский центр.

Bulletin of the Ural Federal District
Security in the Sphere of Information No. 2(60) / 2026
Signed to print 30.06.2026. Date of publication of the 06.07.2026.
Format 70×108 1/16. Screen printing. Conventional printed sheet 5,78. Circulation – 50 issues.
Order 128/245.
Open price.

Printed in the printing house of the Publishing Center of FGAOU VO "SUSU (NIU)".
SUSU, Publishing Center, 76, Lenina Str., Chelyabinsk, 454080