МОДЕЛИ И МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОГО РАСПОЗНАВАНИЯ КИБЕРУГРОЗ КРИТИЧЕСКИ ВАЖНЫМ КОМПЬЮТЕРНЫМ СИСТЕМАМ

АКТУАЛЬНОСТЬ ИССЛЕДОВАНИЙ

- обуславливается:
- важностью КВКС в национальной безопасности и экономике государства.
- необходимостью обеспечить безопасность технологических процесса в КВКС и их информационной составляющей, роль которой постоянно растет.
- потенциальными уязвимостями КВКС, что обусловлено появлением новых классов кибератак, широким распространением беспроводных коммуникаций, систем навигации с использованием GPS, ГЛОНАСС, GALILEO, систем видеонаблюдения (SC), технологий связи GSM, VSAT, систем диспетчерского управления (SCADA, HMI), PLC.
- несовершенством существующих методов киберзащиты, а также изменяющимся характером действий атакующей стороны, при чем в качестве последней могут выступать не только хакеры одиночки или группа хакеров, но и кибервойска стран потенциальных противников, в результате чего состояние КВКС может стать небезопасным.

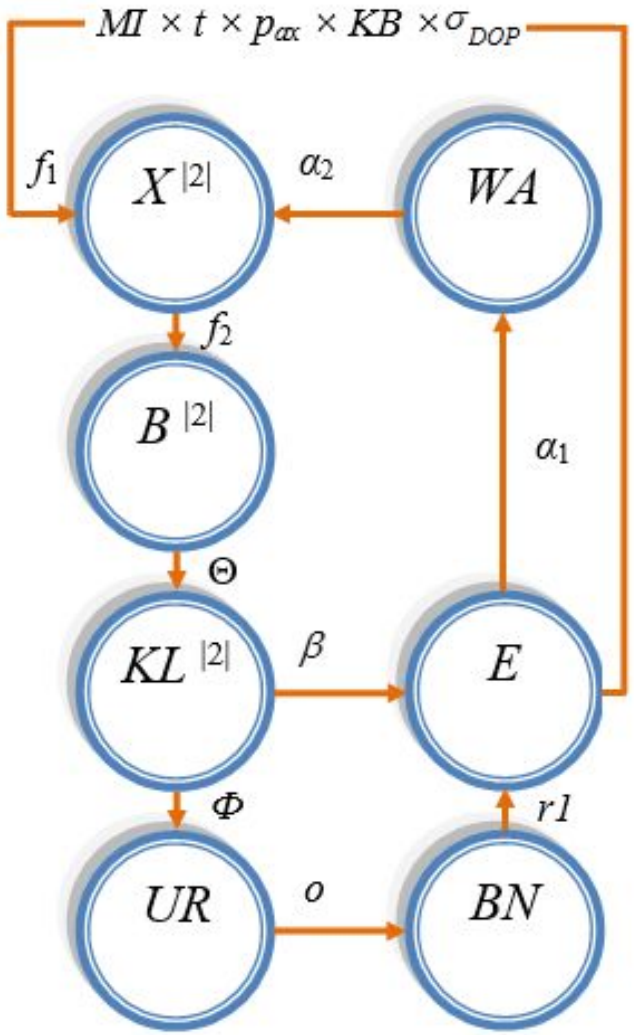
ЦЕЛЬ И ЗАДАЧИ ИССЛЕДОВАНИЯ

- **Целью диссертационной работы является развитие моделей и методов защиты критически важных компьютерных систем на основе интеллектуального распознавания киберугроз в условиях постоянного увеличения количества дестабилизирующих воздействий на конфиденциальность, целостность и доступность информации.**
- Для достижения поставленной цели необходимо решить следующие задачи:
- 1. Разработать метод интеллектуального распознавания угроз, аномалий и кибератак, позволяющий обеспечить кибербезопасность КВКС на основе применения инновационных интеллектуальных систем киберзащиты для повышения устойчивости КВКС к кибератакам.
- 2. Разработать модель интеллектуального распознавания с использованием логических процедур выявления аномалий и кибератак, базирующуюся на покрытиях матриц признаков (МП) и понятии элементарного классификатора (ЭК).
- 3. Минимизировать количество обучающих выборок для признаков, расположенных в репозитории интеллектуальной системы распознавания угроз, аномалий и кибератак.
- 4. Выполнить имитационное моделирование основных компонентов КВКС и подсистемы киберзащиты, основанной на предложенных моделях интеллектуального распознавания угроз, аномалий и кибератак в КВКС.

МОДЕЛЬ РАСПОЗНАВАНИЯ КИБЕРУГРОЗ, КИБЕРАТАК ИЛИ АНОМАЛИЙ В КРИТИЧЕСКИ ВАЖНЫХ КОМПЬЮТЕРНЫХ СИСТЕМАХ

Базисом для построения классификаторов киберугроз для СРКА, могут стать различные данные, например, описательная информация которая представлена в форме двоичного представления трудно объяснимых признаков кибератаки или угрозы для ИБ – $\{p_{ak1}, \ldots, p_{akn}\}$ в КВКС. Также могут быть использованы: диапазоны пороговых значений и параметры входящего и исходящего трафика, непредусмотренные адреса пакетов, атрибутов, временных параметров, запросов и т. д.

КАТЕГОРИЙНАЯ МОДЕЛЬ ИНТЕЛЛЕКТУАЛЬНОГО РАСПОЗНАВАНИЯ АНОМАЛИЙ И КИБЕРАТАК С ИСПОЛЬЗОВАНИЕМ ЛОГИЧЕСКИХ ПРОЦЕДУР, БАЗИРУЮЩУЮСЯ НА ПОКРЫТИЯХ МАТРИЦ ПРИЗНАКОВ



Принято: MI – общее число киберугроз для КВКС; PA – число возможных целей нарушителя в защищаемой КВКС; B_{p_a} – множество номеров киберугроз, реализуемых нарушителем при достижении p_a - й цели, например, в ходе атаки на КВКС.

Этап 1. Квантор $\Theta: B^{|2|} \to KL^{|2|}$ позволяет разбить пространство признаков ОР (аномалий, кибератак или угроз) на классы $KL^{|2|}$ (пара: эталон – ОР). Проверка гипотезы о принадлежности ОИО к классу KL, реализуется оператором Φ.

Этап 2. Точность распознавания контролируется параметром BN (оператором о). Множество UR учитывает количество статистических гипотез при формировании ОИО. Оператор $r1$ формирует множество E для оценки результативности применения ОИО для класса KL.

Этап 3. Оператор β используется в процессе оптимизации системы контрольных отклонений ОИО.

Этап 4. Схема последовательно замыкается множеством WA, включающим операторы $\alpha_1: E \to WA$ и $\alpha_2: WA \to X^{|2|}$, которые отслеживают реализацию ОИО в процессе обучения адаптивной СРКА. t – множество моментов времени, в ходе которых происходит снятие параметров «слепков» ИБ КВКС; p_a – признаки ОР; G_{DOP} – ЭК, используемые в алгоритмах формирования ОИО; KB – база знаний (ОИО, МП и др.) для идентификации ОР; $X^{|2|}$ – матрица эталон; $B^{|2|}$ – бинарная учебная матрица; f_1, f_2 – процедуры формирования матрицы признаков в ее описательном и бинарном представлении, соответственно.

МЕТОД ИНТЕЛЛЕКТУАЛЬНОГО РАСПОЗНАВАНИЯ УГРОЗ, АНОМАЛИЙ И КИБЕРАТАК

- Особенности метода - возможность получить результат даже в ситуации, когда нет данных по функциям, описывающим распределение значений признаков кибератаки (или аномалий).
- В рамках метода предложены логические процедуры распознавания угроз.
- Парадигмой построения ЛПРУ (логических процедур распознавания угроз) является отыскания информативных подописаний (или фрагментов описаний) объектов. Эти фрагменты при создании, конкретных проектных решений для СРКА, позволят однозначно делать вывод о наличии (или отсутствии) атаки (аномалии, угрозы) в рамках класса.
- Исходные данные – признаки аномалий, атак и киберугроз.
- Информативными положим фрагменты, отражающие характерные закономерности при описании объекта, используемого в ходе обучения СРКА. Тогда, наличие (отсутствие) фрагмента(тов) в описании объекта, проходящего классификацию, дает возможность отнести его к определенному классу. В ЛПРУ информативным положим фрагмент(ты), который имеется в описаниях объектов рассматриваемого класса кибератак, но отсутствует в описании объектов других классов.

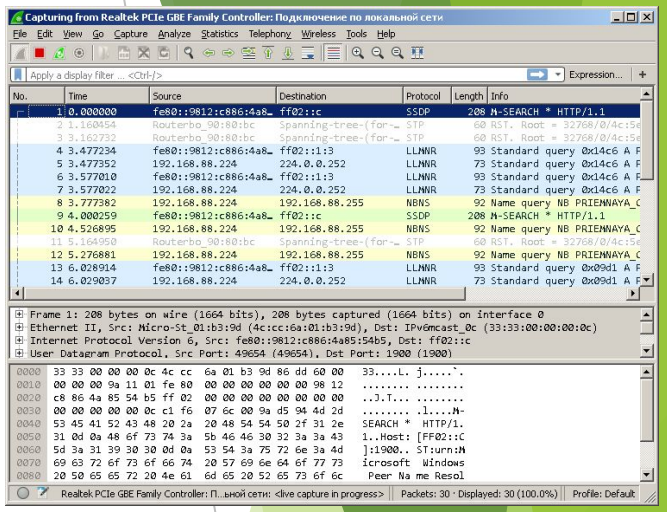
Этапы метода

Этап 1. Для каждого класса кибератак (киберугроз, аномалий, уязвимостей и т.п.) выполняется построение множества ЭК с заранее заданными параметрами.

При этом сделаем следующие допущения для ЭК:

1) используются ЭК, которые присутствуют в описаниях объектов анализируемого в данный момент класса, но их нет в описании объектов других классов;
2) описательный наборов показателей (признаков) задан в двоичной форме (0010101). При этом ОИО характеризуют все объекты данного класса. А, следовательно, ОИО имеют большую информативность.

формировании тестовых выборок ОИО (например, используем Wireshark)

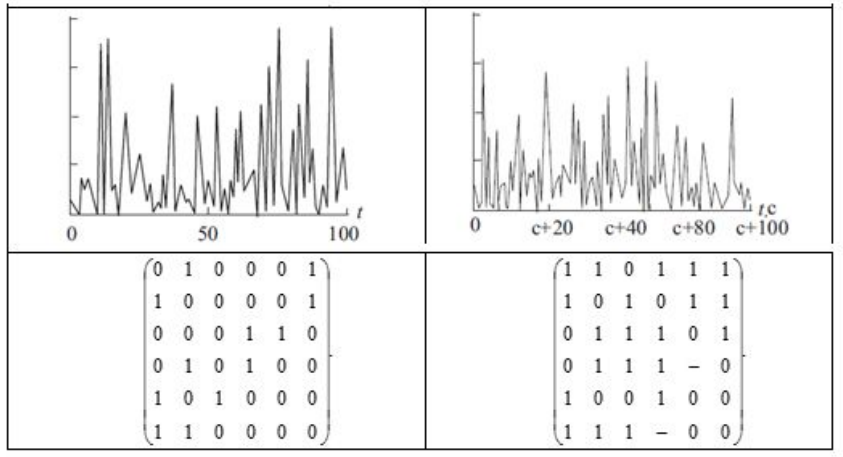


Этап 2. Построение интеллектуальных агентов распознавания киберугроз и множества элементарных классификаторов для моделируемого класса объектов:
Задается характеристическая функция для класса объекта распознавания (**например, для DoS атак**)

$$RTO_{i+1} = RTO_i + 2 * RTO_i;$$
$$RTO_0 = 1 \text{ секунда}.$$

Нет атаки **Есть атака**

Этап 3. Находим дизъюнктивную нормальную форму (ДНФ) (или СДНФ), которая реализует характеристическую функцию.



Этап 4. Находим допустимую (максимальную) конъюнкцию, которая определяет принадлежность объекта к рассматриваемому классу

$$\Re = p_{\alpha x j_1}^{\sigma_{DOP_1}} \ldots p_{\alpha x j_{r_{pa}}}^{\sigma_{DOP_{r_{pa}}}}$$

$$D^*_i = \bigvee_{t \neq \beta_{i1}} p^{\eta}_{\alpha x 1} \vee \bigvee_{t \neq \beta_{i2}} p^{\eta}_{\alpha x 2} \vee \ldots \vee \bigvee_{t \neq \beta_{iMI}} p^{\eta}_{\alpha x MI}, \; i = 1,2,..,u.$$

ИСПОЛЬЗОВАНИЕ ЭЛЕМЕНТАРНЫХ КЛАССИФИКАТОРОВ В МОДЕЛИ

Задача сократить количество обучающих выборок за счет выбора ЭК с высокой информативностью для каждого класса ОР

Близость объекта $sp_{an} = (\alpha p_{a1}, \alpha p_{a2}, \ldots, \alpha p_{aMI})$ из PA и ЭК $\sigma_{DOP} = (\sigma_{DOP_1}, \ldots, \sigma_{DOP_r})$, порожденного набором NP_{p_a}, оценивается так:

$$BN(\sigma_{DOP}, sp_a, NP_{pa}) = \begin{cases} 1, & \text{если } \alpha p_{j_{ti}} = \sigma_{DOP\,ti} \text{ при } ti = 1,2,\ldots,r_{p_a}, \\ 0 & \text{в противном случае.} \end{cases}$$

Каждый алгоритм, используемый для распознавания, в рамках класса KL, $KL \in \{KL_1, \ldots, KL_l\}$, обозначим – AL. Тогда, будем для каждого класса рассматривать подмножество $MC^{AL}(KL)$ множества MC.

Элементарный классификатор (ЭК), используемый в соответствующем алгоритме $\sigma_{DOP} = (\sigma_{DOP_1}, \ldots, \sigma_{DOP_r})$, образованный информационными подописаниями (ИП) из NP_{pa}, должен обладать как минимум одним из свойств: 1) фрагмент группы (sp'_a, NP_{pa}), где $sp'_a \in KL$, тождественны с $\sigma_{DOP} = (\sigma_{DOP_1}, \ldots, \sigma_{DOP_r})$; 2) только часть фрагментов (sp'_a, NP_{pa}), где $sp'_a \in KL$, совпадает с $\sigma_{DOP} = (\sigma_{DOP_1}, \ldots, \sigma_{DOP_r})$; 3) ни один фрагмент группы (sp'_a, NP_{pa}), где $sp'_a \in KL$, не совпадает с $\sigma_{DOP} = (\sigma_{DOP_1}, \ldots, \sigma_{DOP_r})$.

ПАРАМЕТР ОПРЕДЕЛЯЮЩИЙ ПРИНАДЛЕЖНОСТЬ ОР К КЛАССУ

Параметр $\Gamma(sp_a, \mathrm{KL})$ определяет принадлежность объекта sp_{a_i} к классу KL:

$$\Gamma(sp_a, \mathrm{KL}) = \frac{1}{|\mathrm{MC}^{\mathrm{AL}}(KL)|} \cdot \sum_{(sp'_a, NP_{pa}) \in \mathrm{MC}^{\mathrm{AL}}(KL)} \mathrm{vop}_{(sp'_a, NP_{pa})} \cdot (1 - BN(sp_a, sp'_a, NP_{p_a})).$$

Оценку информативности значения показателя sp_{a_i}, т.е. в нашей задаче, киберугрозы или кибератаки p_{axj} можно с помощью следующей величины – $\mathrm{IZ}_{p_{axj}}$:

$$\mathrm{IZ}_{p_{axj}} = \frac{\sum_{(sp'_a, NP_{pa}) \in \mathrm{MC}^{\mathrm{AL}}(KL),\ p_{axj} \in NP_{pa}} \mathrm{vop}_{(sp'_a, NP_{pa})}}{\sum_{\substack{(sp'_a, NP_{pa}) \in \mathrm{MC}^{\mathrm{AL}}(KL) \\ p_{axj} \in NP_{pa}}} \mathrm{vop}_{(sp^*_a, NP^*_{pa})}}.$$

где vop - **функция значимости голосования по представительным наборам за пару «ОИО – ОР» на основе ЛПРУ и ЭК объекта распознавания.**

ПРЕИМУЩЕСТВА ЛПРУ

Преимущества ЛПРУ обеспечиваются:

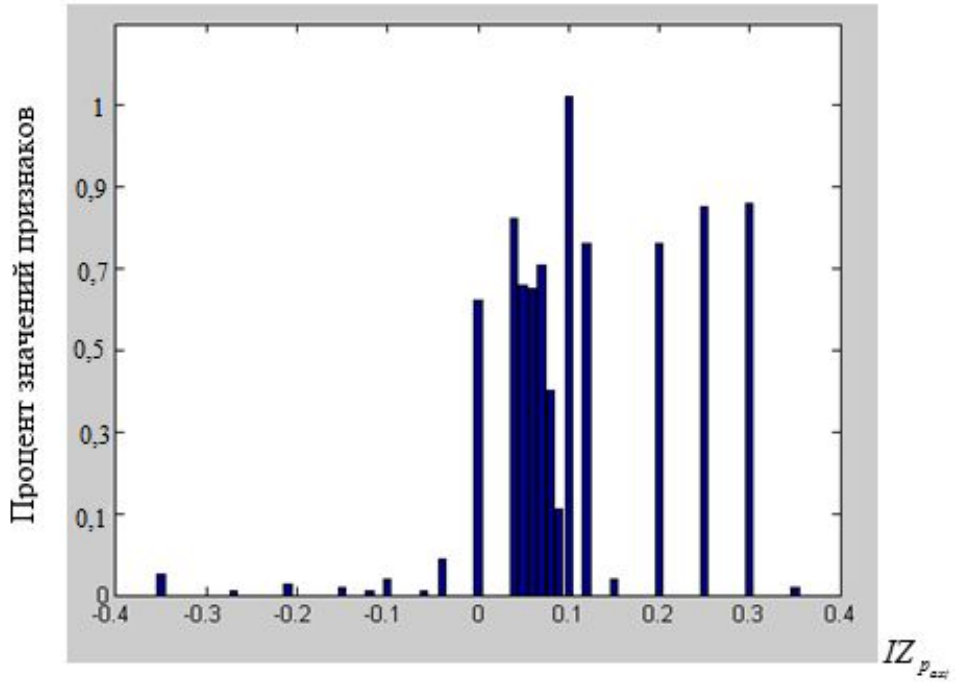
– получением функции классификации ОР с минимальным уровнем ошибки;

– применением линейных ЭК для работы с нелинейными данными;

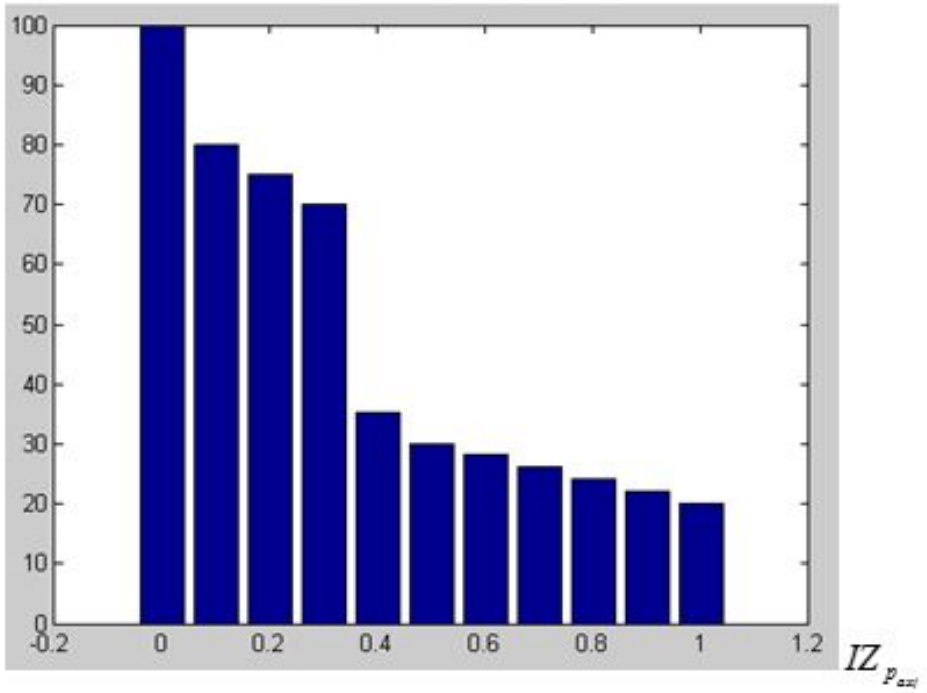
– работой со сложно структурированными данными за счет использования ЭК, описывающих, например, шаблоны кибератак;

– быстрой заменой правил в случае изменения структуры анализируемых данных, без модификации самого алгоритма распознавания ОР.

РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ МЕТОДА И МОДЕЛИ В MATLAB 7



Распределение типичности значений информативности показателей (признаков) для задачи распознавания компьютерной атаки



Распределение информативности признаков для задачи компьютерной DoS атаки

НАУЧНАЯ НОВИЗНА РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ

- *Разработаны:*
- – метод интеллектуального распознавания киберугроз, базирующийся на определении конъюнкций по покрытиям матриц, содержащих бинарные информативные характеристики признаков для классов угроз, аномалий и кибератак, и отличающийся от существующих применением логических функций и нечетких множеств признаков кибератак, что позволяет создавать эффективные аналитические, схемотехнические и программные решение СЗИ для КВКС;
- – модель распознавания и формирования решающего правила для логических процедур идентификации киберугроз и атак, основанная на процедуре анализа критичности отдельных компонентов КВКС, покрытиях матриц бинарных признаков и понятии элементарного классификатора, и в отличие от существующих, обеспечивающая возможность интеллектуального распознавание с минимальным количеством ошибок, а также, учитывать трудно объяснимы признаки аномалий, угроз и кибератак;

СРАВНЕНИЕ ПРЕДЛОЖЕННОГО МЕТОДА И МОДЕЛИ С ДРУГИМИ

Класс объектов для распознавания (кибератак)	Количество признаков (признаки и их информа-тивность)	Среднее количество правил, матриц и шагов для обучения на объект (Правила/Матрицы/Шаги для обучения)		
		Модели и алгоритмы последовательного перебора признаков	Статистические модели прогнозирования состояний	Наша модель Модель, основанная на обучающих выборках и ЭК класса
Сетевые атаки через корпоративную систему	11	200/30/2000	350/65/2000	60/10/2000
Атаки на стандартные компоненты ПО КВКС	19	350/50/3500	450/35/3500	30/15/1500
Сетевая разведка	15	320/40/2500	120/30/2500	70/20/2000
Атаки, направленные на подбор паролей	12	230/15/1500	180/25/1300	25/20/1500
Атаки типа Man-in-the-Middle	9	300/40/4000	350/30/3000	40/20/2000
DoS/DDoS – атаки	9	150/25/2500	170/25/2000	30/15/1500
Вирусные атаки	21	400/50/2700	400/60/2500	35/25/1700
Атаки на ERP системы через протокол HARD	5	170/30/2700	210/50/2300	60/35/1900
Атаки на компоненты ЛВС	9	260/25/2400	200/40/2500	45/35/2000
Атаки систем SCADA	7	600/70/4000	800/60/3000	150/50/3500
Атаки на HMI	3	500/50/3000	400/60/3000	70/30/2600
Атаки подмена узла («атака воронки»)	15	150/35/1500	100/55/1500	30/15/1500
Компрометация узла сбора данных	5	250/30/1700	190/35/1800	30/20/1300
Подмена маршрутизатора	11	300/40/2300	380/60/2500	35/20/1700
Съем информации с периферийных устройств	15	150/25/1500	75/20/1400	45/10/1000
Атаки на системы спутниковой навигации	9	90/30/4000	150/50/4000	20/15/150

ПРАКТИЧЕСКАЯ ЗНАЧИМОСТЬ НАУЧНЫХ РЕЗУЛЬТАТОВ

- Предложенные в диссертации модели и информационные технологии интеллектуального распознавания киберугроз КВКС доведены до практической реализации путем разработки прикладного программного обеспечения, позволяющего повысить эффективность распознавания в зависимости от класса аномалий, кибератак и угроз, до 85–98%, а также на 25–30% уменьшить время отладки проектов системы киберзащиты за счет имитационного моделирования кибератак на модули КВКС, создавать эффективные аналитические, схемотехнические и программные реализации СЗИ.

Опубликовано

- 1. Гулжанат Бекетова, Бахытжан Ахметов, Александр Корченко, Валерий Лахно. РАЗРАБОТКА МОДЕЛИ ИНТЕЛЛЕКТУАЛЬНОГО РАСПОЗНАВАНИЯ АНОМАЛИЙ И КИБЕРАТАК С ИСПОЛЬЗОВАНИЕМ ЛОГИЧЕСКИХ ПРОЦЕДУР, БАЗИРУЮЩУЮСЯ НА ПОКРЫТИЯХ МАТРИЦ ПРИЗНАКОВ . // *Ukrainian Scientific Journal of Information Security, 2016, vol. 22, issue 3, p. 143-149.*
- *2.*