8. Kwok, R. C. W. Collaborative assessment in education: an application of a fuzzy GSS [Text] / R. C. W. Kwok, J. Ma, D. Vogel, D. Zhou // Information & Management. – 2001. – Vol. 39, Issue 3. – P. 243–253. doi: 10.1016/s0378-7206(01)00093-3
9. Проблеми інтеграції національних закладів вищої освіти до Європейського освітнього середовища. Т. 2 [Текст]: матеріали міжнародної наук.-метод. конф.. – Х.: Форт, 2012. – 136 с.
10. Шевченко, В. А. Проверка эффективности обучения студентов с помощью методов непараметрической статистики [Текст] / В. А. Шевченко // Вестник ХНАДУ. – 2013. – Вып. 60. – С. 18–21.
11. Метешкин, К.А. Нечеткое представление результатов кластеризации студентов [Текст] / К. А. Метешкин, В. А. Шевченко // Открытые информационные и компьютерные интегрированные технологии. – 2012. – Вып. 56.– С. 162–168.
12. Шевченко, В. А. Информационная технология формирования индивидуальной траектории самостоятельной работы студентов [Текст] / В. А. Шевченко // Вісник НТУ «ХПІ». – 2015. – № 21 (1130). – С. 76–83.

Запропоновано лінгвістичну модель, що описує основні активи інформаційної системи, які підлягають захисту, і ризики інформаційної безпеки. Побудована нечітка ієрархічна модель, яка містить лінгвістичні змінні і нечіткі бази знань. Дана модель дозволяє дати природну оцінку ризиків, що загрожують активам інформаційної системи

Ключові слова: методологія Coras, актив, ризик, нечіткі бази знань, лінгвістичні змінні

Предложена лингвистическая модель, описывающая основные защищаемые активы информационной системы и риски информационной безопасности. Построена нечеткая иерархическая модель, которая содержит лингвистические переменные и нечеткие базы знаний. Данная модель позволяет дать естественную оценку рисков, угрожающих активам информационной системы

Ключевые слова: методология Coras, актив, риск, нечеткие базы знаний, лингвистические переменные

УДК 004.056
DOI: 10.15587/1729-4061.2015.48239

РАЗРАБОТКА ЛИНГВИСТИЧЕСКОЙ МОДЕЛИ ОЦЕНКИ РИСКОВ АКТИВОВ ИНФОРМАЦИОННОЙ СИСТЕМЫ

В. О. Шапорин
Старший преподаватель*
E-mail: shaporin_v@ukr.net
П. М. Тишин
Кандидат физико-математических наук, доцент*
E-mail: tik88@mail.ru
Р. О. Шапорин
Кандидат технических наук, доцент*
E-mail: shaporin@ukr.net
Н. Б. Копытчук
Доктор технических наук, профессор*
E-mail: knb47@mail.ru
*Кафедра компьютерных интеллектуальных систем и сетей
Одесский национальный политехнический университет
пр. Шевченко, 1, г. Одесса, Украина, 65044

1. Введение

Проектирование информационной безопасности состоит из множества этапов, самыми скрупулезными и длительными из которых являются анализ рисков безопасности и построение политик безопасности. Этап анализа рисков играет крайне важную роль при построении комплекса мер по информационной защите данных. Точность, объективность и компетентность действий команды проектировщиков напрямую влияют на адекватность оценки того, какие активы организации необходимо защитить, какие угрозы и риски угрожают им, и какие меры исправления и предотвращения необходимо применить.

На сегодняшний день процесс анализа рисков информационной безопасности сводится к действиям разработчиков, основанным на личном опыте. Реже используются инструментальные средства анализа, построенные на вероятностных зависимостях процессов. Первый вариант требует длительного обучения и не всегда позволяет объективно рассмотреть конкретную ситуацию, второй вариант требует построения вероятностных зависимостей и функций распределения, что не всегда можно сделать, и не позволяет использовать накопленный опыт разработчика.

В области анализа рисков достаточно давно успешно применяется методология Coras, которая позволяет языком диаграмм описывать процессы, происходящие

в информационной системе, проводить описание активов, угроз и рисков.

2. Анализ литературных данных и постановка задачи

В работе [1] построены вероятностные модели оценки риска осуществления атаки с использованием сетей Петри-Маркова. Данный подход позволяет оценить угрозы в некоторой информационной системе с учетом параметров этой системы. Однако при оценке угроз не применялась методология Coras [2], которая на настоящее время является объединяющей методологии и подходы, такие как Event-Tree-Analysis, цепи Маркова, HazOp и FMECA.

В свою очередь, в работе [3] для оценки процессов, происходящих в информационной системе, использовались марковские цепи. Однако данный подход не вводились параметры системы, как в работе [2].

С другой стороны параметры системы, задаваемые в работе [2], часто нельзя определить точно. В работе [4] рассмотрен многокритериальный подход к обеспечению информационной безопасности, однако к множеству критериев применяются операции свертки, что также приводит к снижению точности анализа самих рисков.

Поэтому для описания таких процессов желательно использовать теорию нечетких множеств [5] и лингвистических переменных [6, 7]. При этом возможно применение методологии Coras, поскольку процессы, происходящие в информационной системе, можно описывать в диаграммах Coras с применением лингвистических переменных.

3. Цель и задачи исследования

Исходя из рассмотренных проблем, целью данного исследования является построение моделей, которые позволяют дать естественную оценку для рассматриваемых процессов, способных хранить и использовать опыт разработчиков и используют язык описания, удобный для восприятия, как проектировщиками, так и владельцами активов.

Для достижения заданной цели необходимо решить следующие задачи:

– необходимо выявить защищаемые активы и определить отношения между элементами диаграммы рисков. Наиболее удобное решение данного этапа – использование методологии Coras.

– построить лингвистические модели рисков для выявленных элементов диаграмм.

– построить нечеткую иерархическую модель оценки рисков, активов и отношений между ними. Для данной модели целесообразно использовать нечеткие лингвистические термы, которые задаются определенным нечетким множеством [8].

4. Разработка моделей для анализа рисков

4. 1. Построение диаграммы рисков

Построение диаграммы заключается в определении ее элементов и отношений между ними. Номенклатура элементов и отношений напрямую зависит от вида диаграммы. В случае построения диаграммы рисков в качестве элементов используются активы, угрозы и риски, мнемоника которых следующая:

– direct asset=da (название, оценка) – прямой актив;

– indirect asset=ia (название, оценка) – непрямой актив;

– risk=r(название, степень риска) – риск;

– human threat deliberate=htd(identifier) – преднамеренная человеческая угроза;

– human threat accidental=hta(identifier) – непреднамеренная человеческая угроза;

– non-human threat=nht(identifier) – нечеловеческая угроза.

Между данными элементами возможны отношения трех типов: инициализация, следствие и H, U, V – определены отношениями (2).

Если один и тот же риск оказывает влияние на оценку нескольких активов, то это считается несколькими рисками.

Следующим этапом является определения отношений между элементами диаграммы. Угрозы приводят к возникновению рисков, и, следовательно, связаны отношением инициализации. Риски могут являться следствием других рисков или влиять на активы системы.

Итоговая диаграмма может быть представлена двумя вариантами – текстовая конструкция, которая описывает элементы и их отношения, либо графическое изображение, удобное для восприятия лиц, нетехнического профиля. Графическое изображение составленной диаграммы рисков Coras представлено на рис. 1.

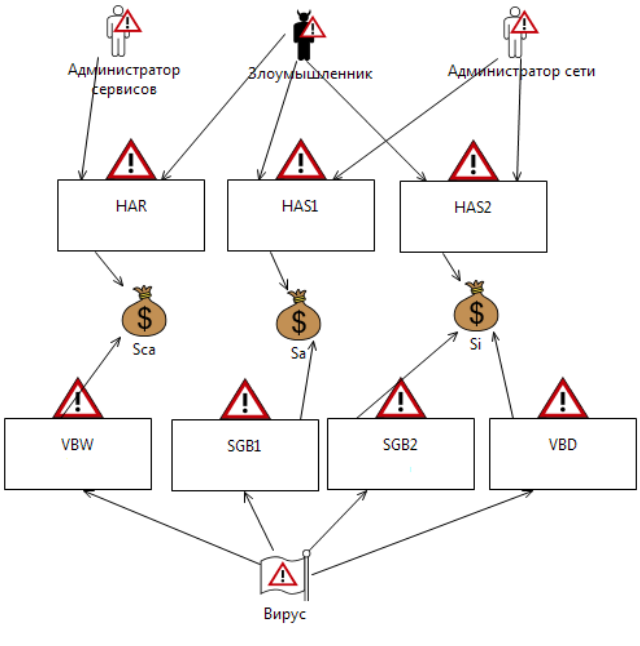

Рис. 1. Диаграмма рисков Coras для серверов и сервисов

Построение данной диаграммы позволяет формально описать поведение рассматриваемых процессов в структурном представлении.

В данной диаграмме используется два типа отношений: C_1 – отношения инициализации риска угрозой, для которых должны быть определены лингвистиче-

ские термы, являющиеся оценками значения вероятности (likelihood); отношения C_{IM} – отношения влияния рисков на оценку актива, для которых должны быть определены лингвистические термы, являющиеся оценками значения следствия (consequence).

4. 2. Построение лингвистической модели риска

В соответствии с целью исследования, модель анализа рисков следует строить с использованием аппарата нечеткой логики и лингвистических переменных в качестве ее параметров. Параметрами модели выступают оценки элементов (активы, риски) и отношения между ними (инициализация, влияние). Следует также отметить, что модель имеет не только входные и выходные параметры, но и внутренние, при этом изменения в оценках данных параметров, также влияют на характеристики модели.

Для формирования входных, выходных и внутренних параметров модели определяются все множества оценок и множества значений отношений.

Оценка важности актива может состоять из количественной оценки (денежный эквивалент, стоимость восстановления и т. п.) или качественной оценки (важный, критичный и т. п.). После определения активов и отношений между ними, оценивание данных активов и определение силы связей отношений производится с помощью лингвистических переменных [9, 10]. Обозначим через T(x) лингвистическую переменную, которая используется при описании актива x. Для каждой лингвистической переменной определено полное ортогональное семантическое пространство [11]. Для этого на множествах значений лингвистических переменных T(x), описывающих систему, определим множества нечетких термов $T_x = \left\{T_x^1, T_x^2, \ldots, T_x^k\right\}$, где k – количество термов введенных для описания T(x).

Как правило, оценка активов в известных методиках анализа рисков определяется денежным эквивалентом стоимости данного актива или стоимостью его восстановления. Однако существуют активы, которые либо не имеют денежного эквивалента стоимости, либо этот эквивалент сложно определить. В связи с этим, в данном исследовании, построение терм-множества оценок активов, определенных соотношениями (1), опирается на выявление степени удовлетворенности данным активом. Данный параметр характеризует состояние актива под воздействием на него сторонних факторов. Для этого сконструируем терм-множество из четырех термов (оценок):

$$T(\text{Asset}) = \left\{\text{низкая, умеренная, средняя, высокая}\right\}.$$

Данное терм-множество задает степени удовлетворенности активом от низкой (ожидается существенный ущерб репутации или финансовому состоянию) до высокой (ожидается несущественный ущерб, которым можно пренебречь).

Таким образом, множество оценок активов информационной системы в модели описывается следующим образом:

$$\text{Asset} = \left\{A_i^k \,|\, i = 1,4;\ k \in \left\{\text{Sa, Si, Sca}\right\}\right\},$$

где i – номер оценки из терма A; k – соответствующий актив.

Риски, определяемые соотношениями (3), характеризуются своей степенью риска и описывают величину опасности от реализации конкретного риска. Сконструированное терм-множество для оценивания степени риска выглядит следующим образом

$$T(\text{Risk}) = \left\{\text{низкая, неопасная, средняя, опасная, высокая}\right\}.$$

Исходя из этого, оценка степени риска будет выглядеть следующим образом

$$\text{Risk} = \{R_i^n \,|\, i = 1,5; k \in \{\text{HAR, HAS1, HAS2, VBW, SGB1, SGB2, VBD}\}\},$$

где i – номер оценки из терма R_i^n, n – соответствующий риск.

Угрозы, которые приводят к возникновению рисков, характеризуются вероятностью их возникновения. В реальных информационных системах данная вероятность обладает высокой степенью неопределенности, что, также как и предыдущие параметры, приводит к необходимости нечеткого лингвистического описания.

$$T(\text{Treat}) = \left\{\text{слабое, среднее, сильное}\right\}.$$

При этом параметр Treat является множеством угроз информационной системы, определенных в (2).

Как было сказано ранее, семантика Coras предусматривает три отношения для диаграммы риска, два из которых присутствуют в данном примере.

Отношение инициализации описывает влияние входных данных на возможность реализации рисков и других сценариев системы. В терминах диаграммы Coras значение данного отношения характеризуется вероятностью. В соответствии с задачами исследования, данной значение рассматривается как степень возможности, что позволяет логически перейти к лингвистическим оценкам. Инициализация тех или иных рисков является процессом, обладающим высокой степенью неопределенности. В связи с этим, при формировании терм множества число термов задано таким образом, чтобы описывать максимальное число возможных ситуаций. Поэтому вводится соответствующее терм-множество оценок:

$C_I = \{$очень низкая, низкая,
ниже среднего, средняя, выше среднего,
высокая, очень высокая$\}$.

Оценка отношения инициализации $C_{I,i}^T$ характеризуется параметрами i – номер оценки из терм-множества, T – элемент инициализатор (угроза).

Оценка влияния описывает степень влияния процессов и рисков на активы. Сконструированный терм представлен следующим образом:

$$C_{IM} = \{\text{слабое, среднее, сильное}\}.$$

Общая оценка отношения $C_{IM,i}^R$ характеризуется параметрами i – номер оценки из терма, R – элемент, источник отношения.

4. 3. Построение нечеткой иерархической модели рисков

Формирование модулей модели производится с помощью объединения пар оценок элементов и пара-

метров отношений, что, по аналогии с построенной диаграммой, позволяет строить модель в виде взаимосвязанных структур. Каждая структура имеет один или более вход и один или более выход. Структура данной модели представлена на рис. 2, *а–в*.

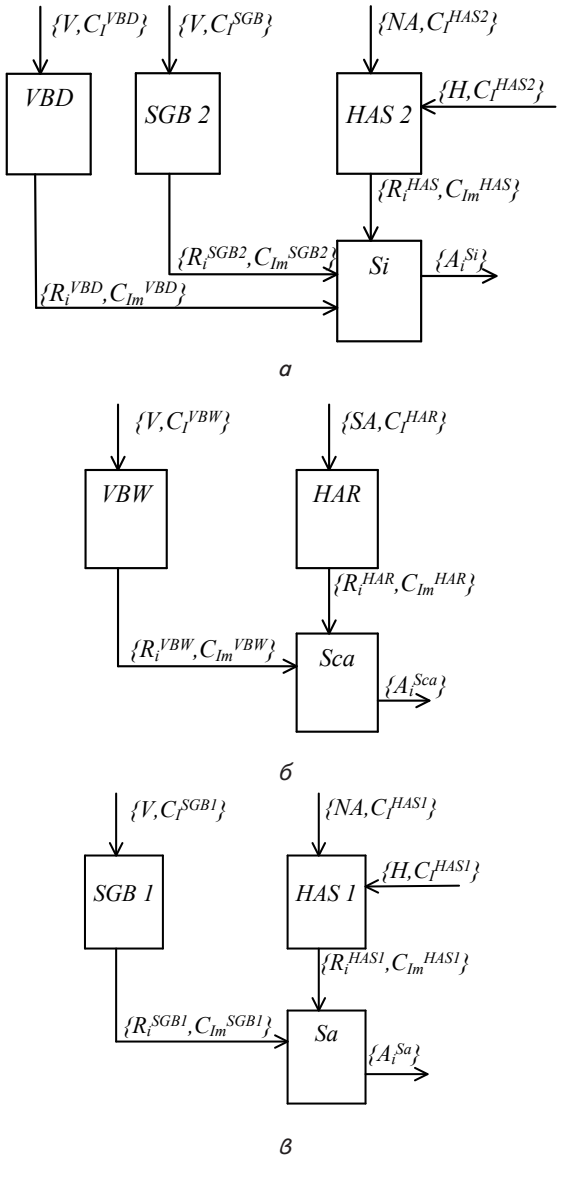

Рис. 2. Иерархическая нечеткая модель анализа рисков: *а* – для актива «целостность сервера»; *б* – для актива «доступность сервисов»; *в* – для актива «доступность сервера»

Разработанная структура представляет собой сеть элементов, взаимосвязанных друг с другом, где, теоретически, уровень связей может достигать полно связной системы. Каждый элемент структуры влияет на как минимум один другой элемент и терпит влияние от других элементов, при этом на каждый отдельный элемент он влияет с различными значениями степеней отношений. Исходя из данного факта, каждый элемент указанной структуры рассматривается как набор баз знаний, которые определяют выходное значение, опираясь на лингвистические значения входных параметров. При построении нечетких баз знаний можно использовать подходы, описанные в [12]. Отметим, что выходное значение базы знаний, также является лингвистическим. Для примера рассмотрим иерархическую модель, представленную на рис. 2, *б*. Модель состоит из двух баз знаний для рисков VBW, HAR и актива Sca. Каждый элемент для рисков имеет два входных параметра и два выходных параметра, элемент для актива имеет четыре входных параметра и один выходной. Каждый выходной параметр формируется отдельной базой знаний и, соответственно, для каждого из рисков имеется две базы знаний. Базы знаний для риска VBW представлены на рис. 3.

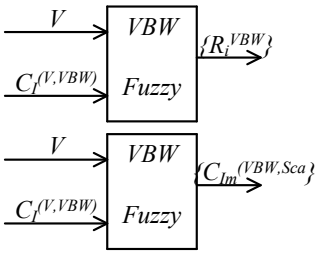

Рис. 3. Базы знаний для риска VBW

Каждая пара входных параметров V и $C_i^{(V,VBW)}$ формирует выходные параметры степени риска R_i^{VBW} и отношения влияния $C_{IM}^{(VBW,Sca)}$ на актив системы. Соответствующие базы знаний, представлены в табл. 1.

Таблица 1

Нечеткие базы знаний для параметров R_i^{VBW} и $C_{IM}^{(VBW,Sca)}$

V	$C_i^{(V,VBW)}$	R_i^{VBW}	$C_{IM}^{(VBW,Sca)}$
Низкая	Очень низкая	Низкая	Очень низкая
Средняя	Очень низкая	Низкая	Очень низкая
Высокая	Очень низкая	Неопасная	Низкая
Низкая	Низкая	Низкая	Очень низкая
Средняя	Низкая	Неопасная	Низкая
Высокая	Низкая	Средняя	Низкая
Низкая	Ниже среднего	Неопасная	Ниже среднего
Средняя	Ниже среднего	Неопасная	Низкая
Высокая	Ниже среднего	Средняя	Ниже среднего
Низкая	Средняя	Неопасная	Средняя
Средняя	Средняя	Средняя	Средняя
Высокая	Средняя	Средняя	Средняя
Низкая	Выше среднего	Средняя	Средняя
Средняя	Выше среднего	Опасная	Выше среднего
Высокая	Выше среднего	Высокая	Выше среднего
Низкая	Высокая	Опасная	Выше среднего
Средняя	Высокая	Высокая	Высокая
Высокая	Высокая	Высокая	Высокая
Низкая	Очень высокая	Опасная	Высокая
Средняя	Очень высокая	Высокая	Очень высокая
Высокая	Очень высокая	Высокая	Очень высокая

Базы знаний для риска HAR представлены на рис. 4.

Каждая пара входных параметров SA и $C_i^{(SA,HAR)}$ формирует выходные параметры степени риска R_i^{HAR} и отношения влияния $C_{IM}^{(HAR,Sca)}$ на актив системы. Соответствующие базы знаний, представлены в табл. 2.

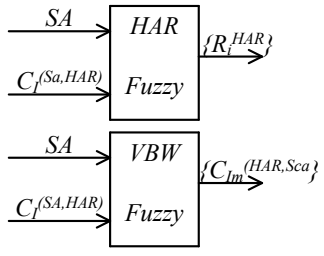

Рис. 4. Базы знаний для риска HAR

Таблица 2

Нечеткие базы знаний для параметров R_i^{HAR} и $C_{IM}^{(HAR,Sca)}$

SA	$C_i^{(SA,HAR)}$	R_i^{HAR}	$C_{IM}^{(HAR,Sca)}$
Низкая	Очень низкая	Низкая	Очень низкая
Средняя	Очень низкая	Низкая	Очень низкая
Высокая	Очень низкая	Неопасная	Низкая
Низкая	Низкая	Низкая	Очень низкая
Средняя	Низкая	Низкая	Очень низкая
Высокая	Низкая	Неопасная	Низкая
Низкая	Ниже среднего	Неопасная	Низкая
Средняя	Ниже среднего	Неопасная	Низкая
Высокая	Ниже среднего	Средняя	Ниже среднего
Низкая	Средняя	Неопасная	Ниже среднего
Средняя	Средняя	Средняя	Средняя
Высокая	Средняя	Средняя	Средняя
Низкая	Выше среднего	Неопасная	Средняя
Средняя	Выше среднего	Опасная	Средняя
Высокая	Выше среднего	Высокая	Выше среднего
Низкая	Высокая	Опасная	Выше среднего
Средняя	Высокая	Опасная	Выше среднего
Высокая	Высокая	Высокая	Высокая
Низкая	Очень высокая	Опасная	Высокая
Средняя	Очень высокая	Опасная	Высокая
Высокая	Очень высокая	Высокая	Очень высокая

Формирование выходных множеств параметров баз знаний можно осуществлять тремя способами:

– задание специальных правил нечеткого вывода;

– на основе статистических данных (обучение);

– на основе экспертных оценок.

База знаний, оценивающая актив системы представлена на рис. 5

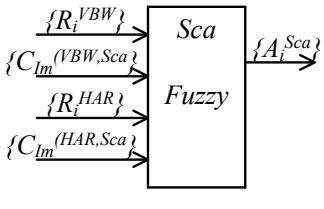

Рис. 5. База знаний для оценки актива Sca

Четверка входных параметров R_i^{VBW}, $C_{IM}^{(VBW,Sca)}$, R_i^{HAR} и $C_{IM}^{(HAR,Sca)}$, полученных на предыдущем уровне иерархии, формируют единственный выходной параметр оценки актива Sca. Фрагмент соответствующей базы знаний, когда выходной параметр описывается термом высокой степени удовлетворенности, представлена в табл. 3.

Таблица 3

Нечеткая база знаний для параметра A_i^{Sca} при i =4

R_i^{VBW}	$C_{IM}^{(VBW,Sca)}$	R_i^{HAR}	$C_{IM}^{(HAR,Sca)}$	A_i^{Sca}
Низкая	Очень низкая	Низкая	Очень низкая	Высокая
Неопасная	Низкая	Неопасная	Низкая	Высокая
Низкая	Очень низкая	Низкая	Очень низкая	Высокая
Средняя	Низкая	Неопасная	Низкая	Высокая
Неопасная	Ниже ср.	Неопасная	Низкая	Высокая
Неопасная	Низкая	Неопасная	Низкая	Высокая
Средняя	Ниже ср.	Средняя	Ниже ср.	Высокая
Неопасная	Средняя	Неопасная	Ниже ср.	Высокая

Таким образом, зная значения лингвистических оценок входных параметров V, $C_i^{(V,VBW)}$, SA и $C_i^{(SA,HAR)}$ и модули нечетких баз знаний, построенные для описанных элементов, имеется возможность осуществить нечеткий вывод, который характеризует выходной нечеткий параметр A_i^{Sca}. В общем случае, разработанная модель представляется нечеткой иерархической системой, с базами знаний нечеткого вывода.

Аналогичный подход применяется для оценки активов «целостность сервера» и «доступность сервера».

Предлагаемый подход позволяет сформировать множество оценок, характеризующих активы информационной системы в зависимости от состояний, в которых данная система находится. Это обобщает подходы, основанные на нечетких когнитивных моделях [13], где оценки рисков или активов сводятся к определению такого параметра, как степень влияния от угроз.

Также, разработанные модели существенно расширяют возможности методологии Coras и позволяют описывать слабо структурированную предметную область, в которой необходимо сформировать множество альтернативных решений, ставя целью поддержку активов на актуальном уровне безопасности.

5. Выводы

Построены лингвистические модели, которые позволяют дать естественную оценку рисков угрожающих активам информационной системы. Особенность рассматриваемого подхода заключается в том, что нечеткие базы знаний, которые используются для построения модели, способны хранить и использовать опыт экспертов в данной проблемной области.

В ходе исследования было достигнуто следующее:

– построена диаграмма рисков. Удалось описать отношение и основные элементы, которые связывают активы ИС, риски и угрозы;

– построены лингвистические модели угроз, рисков и активов – данные модели позволяют описывать множества активов и рисков в случае нечетко заданных или качественно описанных параметров;

– построена нечеткая иерархическая модель анализа рисков. Модель формируется из множества

модулей. Данные модули составляют иерархические цепочки, демонстрирующие влияние на конкретный актив. Каждый модуль является нечеткой базой знаний, которая принимает нечеткое множество параметров, обрабатывает их, и передает новое нечеткое множество на следующий уровень. Это позволяет оценивать влияние угроз и других факторов на активы ИС.

Процесс анализа рисков, при использовании данного подхода, сводится к комбинированию уже имеющихся блоков иерархической нечеткой модели или, при необходимости, формировании собственных.

Литература

1. Миронова, В. Г. Сети Петри–Маркова как инструмент создания аналитических моделей для основных видов несанкционированного доступа в информационной системе [Текст] / В. Г. Миронова, А. А. Шелупанов, М. А. Сопов // Доклады ТУСУРа. – 2012. – № 1 (25), часть 2. – С. 20–24.
2. Lund, M. S. Model-Driven Risk Analysis [Text] // M. S. Lund, B. Solhaug, K. Stolen // Springer-Verlag. Berlin, 2011. – P. 55–62. doi: 10.1007/978-3-642-12323-8
3. Yaqub, S. C. Relating CORAS diagrams and Markov chains [Text]: master thesis / S. C. Yaqub. – Shahbaz Chaudhary Yagub [University of Oslo], 2007.
4. Zadeh, L. A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic [Text] / L. A. Zadeh // Fuzzy Sets and Systems. – 1997. – Vol. 90, Issue 2. – P. 111–127. doi: 10.1016/s0165-0114(97)00077-8
5. Ажмухамедов, И. М. Моделирование на основе экспертных суждений процесса оценки информационной безопасности [Текст] / И. М. Ажмухамедов // Вестник АГТУ. Сер. Управление, вычислительная техника и информатика. – 2009. – № 2. – С. 101–109.
6. Nieto-Morote, A. A fuzzy approach to construction project risk assessment [Text] / A. Nieto-Morote, F. Ruz-Vila // International Journal of Project Management. – 2011. – Vol. 29, Issue 2. – P. 220–231. doi: 10.1016/j.ijproman.2010.02.002
7. Шапорин, В. О. Разработка нечетких лингвистических моделей атак для анализа рисков в распределенных информационных системах [Текст]: 15-я межд. науч.-прак. конф. / В. О. Шапорин, П. М. Тишин, Н. Б. Копытчук, Р. О. Шапорин // Современные информационные и электронные технологии. – Одесса, 2014. – С. 131–132.
8. Шапорин, В. О. Оценка вероятности проведения атаки на сетевые ресурсы с использованием аппарата нечеткой логики [Текст] / В. О. Шапорин П. М. Тишин, Н. Б. Копытчук, Р. О. Шапорин // Электротехнические и компьютерные системы. – 2013. – № 12 (88). – С. 95–101.
9. Нестеренко, С. А. Модель онтологии априорного подхода прогнозирования проблемных ситуаций в сложных вычислительных системах [Текст] / С. А. Нестеренко, П. М. Тишин, А. С. Маковецкий // Электротехнические и компьютерные системы. – 2013. – № 10 (86). – С. 111–119.
10. Копытчук, Н. Б. Процедура создания нечетких моделей анализа рисков в сложных вычислительных системах [Текст] / Н. Б. Копытчук, П. М. Тишин, М. В. Цюрупа // Электротехнические и компьютерные системы. – 2014. – № 13(89). – С. 215–222.
11. Рыжов, А. П. Элементы теории нечетких множеств и ее приложений [Текст] / А. П. Рыжов. – М.: Диалог-МГУ, 2003. – С. 53–65.
12. Штовба, С. Д. Проектирование нечетких систем средствами MATLAB [Текст] / С. Д Штовба. – М.: Горячая линия – Телеком, 2003. – С. 263–275.
13. Ажмухамедов, И. М. Решение задач обеспечения информационной безопасности на основе системного анализа и нечеткого когнитивного моделирования [Текст]: монография / И. М. Ажмухамедов. – Астрахань, 2012. – 344 с.