Перевод статьи:
Glushenko S.A. An adaptive neuro-fuzzy inference system for assessment of risks to an organization's information security
Business Informatics. 2017. No. 1 (39). P. 68–77.

Адаптивная нейро-нечеткая система оценки рисков информационной безопасности организации[1]

С.А. Глушенко
кандидат экономических наук,
старший преподаватель кафедры информационных систем и прикладной информатики
Ростовский государственный экономический университет (РИНХ)
Адрес: 344002, г. Ростов-на-Дону, ул. Большая Садовая, д. 69
E-mail: gs-gears@yandex.ru

Аннотация

В статье обосновывается важность применения оценки рисков при реализации системы обеспечения информационной безопасности. Рассматриваются наиболее распространенные методики оценки риска и предлагается использовать для этих целей теорию нечеткой логики. Описывается предложенная нечеткая продукционная модель (НПМ), в которой определены семь входных лингвистических переменных, характеризующих факторы риска, четыре выходных лингвистических переменных, характеризующих риски различных областей информационной безопасности, а также четыре базы правил.

Отмечается, что НПМ является первым приближением для рассматриваемой предметной области и требует оптимизации с целью минимизации ошибки выходов модели. Рассматриваются наиболее распространенные методы оптимизации параметров нечетких моделей и обосновываются преимущества применения методов, основанных на использовании нейро-нечетких сетей (ННС).

Описывается процесс преобразования элементов нечеткой модели, таких как блок фаззификации, блок базы правил и блок дефаззификации во фрагменты нейронной сети. Результатом данного процесса является нейро-нечеткая сеть, соответствующая нечеткой модели.

Построение разработанной ННС осуществляется на основе системы нейро-нечеткого вывода (adaptive neuro-fuzzy inference system, ANFIS) посредством применения специализированного пакета Neuro-Fuzzy Designer программного средства MATLAB. Обучение модели было выполнено гибридным методом, который представляет собой комбинацию методов наименьших квадратов и обратного распространения ошибки. Результатом данного процесса является оптимизация (настройка) параметров функций принадлежности входных лингвистических переменных.

Использованный подход нейро-нечеткого моделирования позволил получить более адекватную нечеткую продукционную модель, которая позволяет проводить лингвистический анализ рисков информационной безопасности организации. Полученные с ее помощью сведения позволяют ИТ-менеджерам определять приоритеты рисков и разрабатывать эффективные планы мероприятий по снижению влияния наиболее опасных угроз.

Ключевые слова: риск, информационная безопасность, лингвистическая переменная, функция принадлежности, нейро-нечеткая сеть, система нейро-нечеткого вывода, дизайнер нейро-нечетких сетей.

Цитирование: Glushenko S.A. An adaptive neuro-fuzzy inference system for assessment of risks to an organization's information security // Business Informatics. 2017. No. 1 (39). P. 68–77. DOI: 10.17323/1998-0663.2017.1.68.77.

[1] Исследование выполнено при финансовой поддержке РФФИ, в рамках научного проекта № 16-31-00285 «Методы и модели нечеткой логики в системах принятия решений управления рисками»

Перевод статьи:
Glushenko S.A. An adaptive neuro-fuzzy inference system for assessment of risks to an organization's information security
Business Informatics. 2017. No. 1 (39). P. 68–77.

Введение

Внедрение информационных технологий и вычислительных средств в производство и управление современными предприятиями является эффективным инструментом, способствующим повышению производительности труда. Однако ИТ-инфраструктура предприятия зачастую приобретает неструктурированный характер, что приводит к неконтролируемому росту уязвимостей и рисков информационной безопасности (ИБ) предприятия в целом.

Информационная безопасность – это «защищенность информации и поддерживающей инфраструктуры от случайных или преднамеренных воздействий естественного или искусственного характера, которые могут нанести неприемлемый ущерб» [1].

В работе [2] выполнен анализ наиболее распространенных методик оценки риска ИБ – NIST [3] и CRAMM [4], описаны их недостатки и предложено использовать для этих целей нечеткую логику. Предложенная нечеткая продукционная модель (НПМ) включает семь входных лингвистических переменных (*таблица 1*), характеризующих факторы риска, четыре выходных лингвистических переменных (*таблица 2*), характеризующих риски различных областей информационной безопасности, а также четыре базы правил (*таблица 3*) [2, 5].

При формировании входных лингвистических переменных могут быть использованы следующие терм-множества, которые определяют уровни факторов [6]:

- Т2 = {Низкий (Н); Высокий (В)};
- Т3 = {Низкий (Н); Средний (С); Высокий (В)};
- Т4 = {Очень Низкий (ОчН); Низкий (Н); Средний (С); Высокий (В)};
- Т5 = {Очень Низкий (ОчН); Низкий (Н); Средний (С); Высокий (В); Очень Высокий (ОчВ)}.

Таблица 1.

Факторы риска информационной безопасности организации

Обозначение	Наименование лингвистической переменной	Вид терм-множества и интерпретация уровней факторов
x_1	Программно-аппаратный уровень защиты	Т3. Н – удовлетворительная, для обеспечения начального уровня защиты; С – достаточная, для базовой информационной защиты; В – полностью соответствует уровню конфиденциальности информации
x_2	Уровень организационной защиты	Т3. Н – слабое планирование и отсутствие мониторинга уязвимостей; С – планирование и мониторинг уязвимостей проводятся нерегулярно; В – своевременное планирование и мониторинг уязвимостей
x_3	Уровень правовой защиты	Т3. Н – обрывочная и неполная документация; С – документация имеется, но недостаточно детальная; В – документация полная и синхронизированная
x_4	Мотивация источника угроз (ИУ)	Т5. ОчН – отсутствует; Н – редкое проявление заинтересованности; С – вполне может заинтересовать; В – скорее всего, заинтересуется; ОчВ – обязательно заинтересуется
x_5	Возможности источника угроз (ИУ)	Т5. ОчН – не обладает; Н – незначительный уровень оснащенности ИУ; С – средний уровень оснащенности; В – достаточно высокий уровень оснащенности; ОчВ – ИУ обладает значительными возможностями
x_6	Рыночная ценность информационного ресурса (ИР)	Т5. ОчН – открытая информация; Н – ИР обладает незначительной ценностью; С – ИР представляет коммерческую тайну; В – высококонфиденциальные данные; ОчВ – катастрофическая ценность для организации (уровень стратегического планирования)
x_7	Объем данных информационного ресурса (ИР) организации	Т5. ОчН – крайне малая часть; Н – меньшая часть; С – половина ИР; В – большая часть; ОчВ – полный объем ИР

Таблица 2.

Показатели риска информационной безопасности организации

Обозначение	Наименование лингвистической переменной	Примечание
y_1	Риск снижения эффективности защиты	Характеризует потенциальную возможность снижения / увеличения эффективности защиты по отношению к требуемой эффективности для конкретного предприятия
y_2	Риск возникновения потенциальных угроз	Характеризует возможность возникновения потенциальных угроз для предприятия
y_3	Риск материального ущерба	Характеризует возможность возникновения материального ущерба для предприятия при нарушениях параметров информационной безопасности предприятия
y_4	Риск ИБ организации	Интегральный риск, характеризующий обеспечение информационной безопасности предприятия

Таблица 3.

Нечеткие продукционные правила модели (фрагмент)

Обозначение правила	Антецедент	Консеквент
	База правил П1	
П1.1	$(x_1 = \text{Н} \wedge x_2 = \text{Н} \wedge x_3 = \text{Н}) \vee (x_1 = \text{С} \wedge x_2 = \text{Н} \wedge x_3 = \text{Н}) \vee (x_1 = \text{Н} \wedge x_2 = \text{С} \wedge x_3 = \text{Н})$	$y_1 = \text{ОчВОР}$
П1.2	$(x_1 = \text{В} \wedge x_2 = \text{Н} \wedge x_3 = \text{Н}) \vee (x_1 = \text{С} \wedge x_2 = \text{С} \wedge x_3 = \text{Н}) \vee (x_1 = \text{Н} \wedge x_2 = \text{В} \wedge x_3 = \text{Н}) \vee (x_1 = \text{С} \wedge x_2 = \text{В} \wedge x_3 = \text{Н}) \vee (x_1 = \text{Н} \wedge x_2 = \text{Н} \wedge x_3 = \text{С}) \vee (x_1 = \text{Н} \wedge x_2 = \text{С} \wedge x_3 = \text{С}) \vee (x_1 = \text{Н} \wedge x_2 = \text{В} \wedge x_3 = \text{С}) \vee (x_1 = \text{Н} \wedge x_2 = \text{Н} \wedge x_3 = \text{В})$	$y_1 = \text{ВОР}$
П1.3	$(x_1 = \text{В} \wedge x_2 = \text{С} \wedge x_3 = \text{Н}) \vee (x_1 = \text{В} \wedge x_2 = \text{В} \wedge x_3 = \text{Н}) \vee (x_1 = \text{С} \wedge x_2 = \text{Н} \wedge x_3 = \text{С}) \vee (x_1 = \text{В} \wedge x_2 = \text{Н} \wedge x_3 = \text{С}) \vee (x_1 = \text{С} \wedge x_2 = \text{С} \wedge x_3 = \text{С}) \vee (x_1 = \text{С} \wedge x_2 = \text{В} \wedge x_3 = \text{С}) \vee (x_1 = \text{С} \wedge x_2 = \text{Н} \wedge x_3 = \text{В}) \vee (x_1 = \text{Н} \wedge x_2 = \text{С} \wedge x_3 = \text{В}) \vee (x_1 = \text{С} \wedge x_2 = \text{С} \wedge x_3 = \text{В}) \vee (x_1 = \text{Н} \wedge x_2 = \text{В} \wedge x_3 = \text{В})$	$y_1 = \text{СОР}$
П1.4	$(x_1 = \text{В} \wedge x_2 = \text{С} \wedge x_3 = \text{С}) \vee (x_1 = \text{В} \wedge x_2 = \text{В} \wedge x_3 = \text{С}) \vee (x_1 = \text{В} \wedge x_2 = \text{Н} \wedge x_3 = \text{В}) \vee (x_1 = \text{В} \wedge x_2 = \text{С} \wedge x_3 = \text{В}) \vee (x_1 = \text{С} \wedge x_2 = \text{В} \wedge x_3 = \text{В})$	$y_1 = \text{НОР}$
П1.5	$x_1 = \text{В} \wedge x_2 = \text{В} \wedge x_3 = \text{В}$	$y_1 = \text{ОчНОР}$

При формировании выходных лингвистических переменных могут быть использованы следующее терм-множества, которые определяют показатели риска [6]:

– Т1 = {Низкая очевидность риска (НОР); Средняя очевидность риска (СОР); Высокая очевидность риска (ВОР)};

– Т2 = {Очень низкая очевидность риска (ОчНОР); Низкая очевидность риска (НОР); Средняя очевидность риска (СОР); Высокая очевидность риска (ВОР); Очень высокая очевидность риска (ОчВОР)}.

Чтобы построить нечеткую модель, необходимо определить все ее элементы: базу правил, число и тип функций принадлежности для каждой переменной модели, параметры функций принадлежности, логические операторы и т. п.

Структура нечеткой продукционной модели оценки рисков ИБ организации приведена на *рисунке 1.*

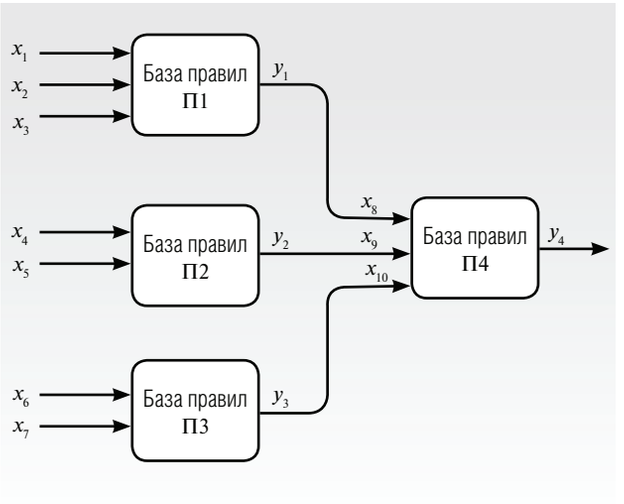

Рис. 1. Структура нечеткой продукционной модели

1. Постановка задачи

Построенная модель основана на экспертных знаниях о моделируемой системе обеспечения информационной безопасности (СОИБ) [2, 5, 7]. Получение информации о системе проводилось с привлечением эксперта предметной области, после чего выполнялось преобразование полученной информации в нечеткую модель. Такой метод является эффективным в том случае, если эксперт владеет всей полнотой знаний о СОИБ. На практике знания экспертов часто являются недостаточно полными и точными, а иногда даже содержат противоречия. Следовательно, необходимо, чтобы модель была основана на объективной информации о системе, в качестве которой могут выступать данные о результатах измерений значений входов и выходов системы [8].

Данные обстоятельства предопределяют актуальность разработки нечеткой самонастраивающейся модели оценки рисков СОИБ. Под настройкой нечеткой модели, прежде всего, понимается процесс определения параметров функций принадлежности входных и выходных лингвистических переменных, при которых минимизируется ошибка выходов модели по отношению к наблюдаемой моделируемой системе.

Для настройки модели, т.е. оптимизации ее параметров, чаще всего применяются следующие методы [9]:

✦ методы, основанные на использовании нейро-нечетких сетей;

✦ поисковые методы;

✦ методы, основанные на кластеризации.

Методы первой группы связаны с преобразованием нечеткой модели в нейро-нечеткую сеть (ННС) и применением для настройки параметров модели методов обучения сети, основанных на измерениях входных и выходных данных системы.

Методы второй группы представляют собой методы прямого поиска оптимальных параметров нечеткой модели. Процесс поиска может быть как упорядоченным, так и неупорядоченным (метод проб и ошибок). Наиболее часто используемым методом упорядоченного поиска является метод, основанный на применении генетических алгоритмов.

Методы, основанные на кластеризации, сочетают настройку параметров модели и ее структуризацию. Они применяются для построения самоорганизующихся нечетких моделей, которые самостоятельно определяют свои существенные входные параметры, задают оптимальное число нечетких множеств для входных и выходных лингвистических переменных и устанавливают форму и число правил.

В настоящее время наиболее изученными являются методы первой группы, которые позволяют [10]:

✧ обеспечить возможность оптимизации (настройки) параметров функций принадлежности лингвистических переменных на основе измерений входных и выходных зависимостей реальной системы;

✧ корректировать нечеткие модели, которые недостаточно точно сформированы экспертами;

✧ расширять формируемые экспертами нечеткие модели на области исследуемой системы, о которых знания экспертов ограничены.

Перечисленные особенности объясняют целесообразность применения методов, основанных на использовании нейро-нечетких сетей, для настройки нечеткой модели оценки рисков информационной безопасности организации.

2. Преобразование нечеткой модели в нейро-нечеткую сеть

Преобразование элементов блока фаззификации представлено на *рисунке 2*, который иллюстрирует преобразование кусочно-линейных функций принадлежности во фрагмент нейронной сети.

Для настройки параметров a_i функций принадлежности в процессе обучения сети необходимо вычислить производные выходных значений блока фаззификации по соответствующим параметрам.

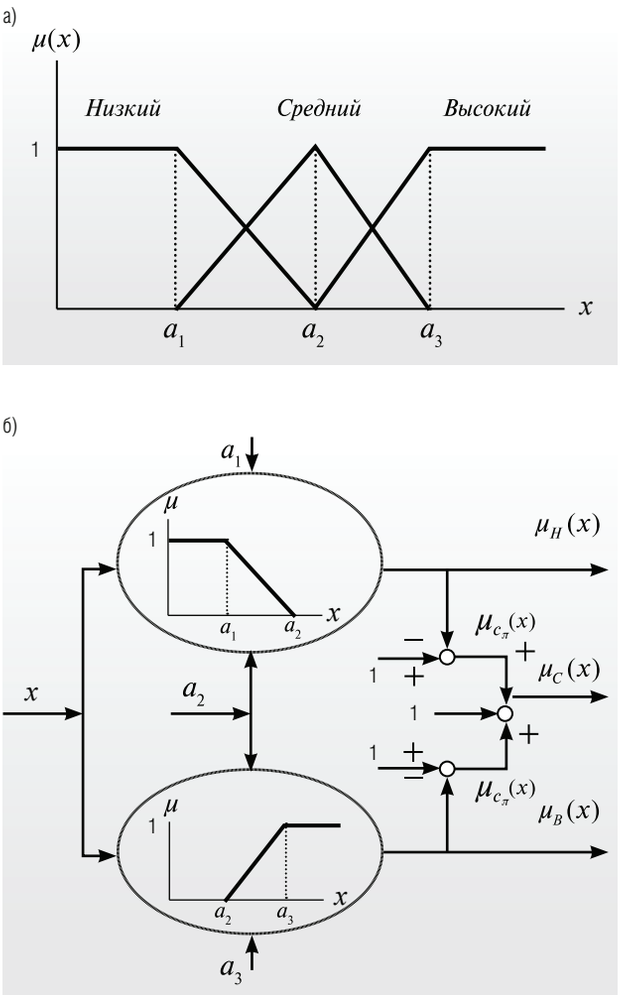

Рис. 2. Преобразование кусочно–линейных функций принадлежности (а) во фрагменты нейронной сети (б)

Результатом блока фаззификации являются вычисленные значения степени принадлежности входных значений нечетким множествам A_{ij}, каждое из которых представляет свою лингвистическую область определения.

Преобразование элементов блока базы правил предполагает представление условия правила в форме фрагмента нейронной сети, при этом выполнение операций «И» и «ИЛИ» может осуществляться с использованием Т-нормы и S-нормы, либо при помощи других операторов.

Входными параметрами блока дефаззификации являются степени активизации $\mu_{B_l}(y)$ нечетких множеств B_l на выходе модели, для преобразования которых в четкое число применяется метод центра тяжести.

В результате нейро-нечеткая сеть, соответствующая нечеткой модели из *таблицы 3*, будет иметь структуру, которая приведена на *рисунке 3*.

3. Применение пакета Neuro-Fuzzy Designer для построения ННС

Построение разработанной ННС проводится на основе системы нейро-нечеткого вывода ANFIS (adaptive neuro-fuzzy inference system) [8, 9], посредством применения специализированного пакета Neuro-Fuzzy Designer программного средства MATLAB [11]. ANFIS представляет собой нейронную сеть с несколькими входами и одним выходом, которые, в свою очередь, являются нечеткими лингвистическими переменными. При этом термы входных и выходных лингвистических переменных описываются функциями принадлежности, что согласуется с разработанной нечеткой самонастраивающейся моделью оценки рисков ИБ.

На этапе фаззификации были заданы треугольные функции принадлежности (*рисунок 4*) для терм-множеств входных (x_1, x_2, x_3) и выходной (y_1) лингвистических переменных (ЛП):

x_1 – ЛП «Программно-аппаратный уровень защиты» (*ПрАпЗащ*);

x_2 – ЛП «Уровень организационной защиты» (*ОргЗащ*);

x_3 – ЛП «Уровень правовой защиты» (*ПравЗащ*);

y_1 – ЛП «Риск снижения эффективности защиты» (*РискЗащ*).

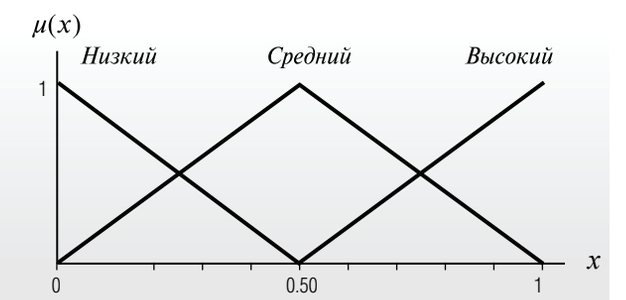

Рис. 4. Функции принадлежности для входной переменной *ПрАпЗащ*

Сгенерированная система нечеткого вывода, которая содержит 27 правил нечетких продукций, приведена на *рисунке 5*.

Обучение ННС было выполнено на основе обучающей выборки, которая содержала 200 наборов, представляющих собой вектор из значений уровней факторов, оказывающих влияние на риск, (входных ЛП) и значений уровня риска ИБ (выходной ЛП). Данные были получены в результате обобщения мнений экспертов предметной области пу-

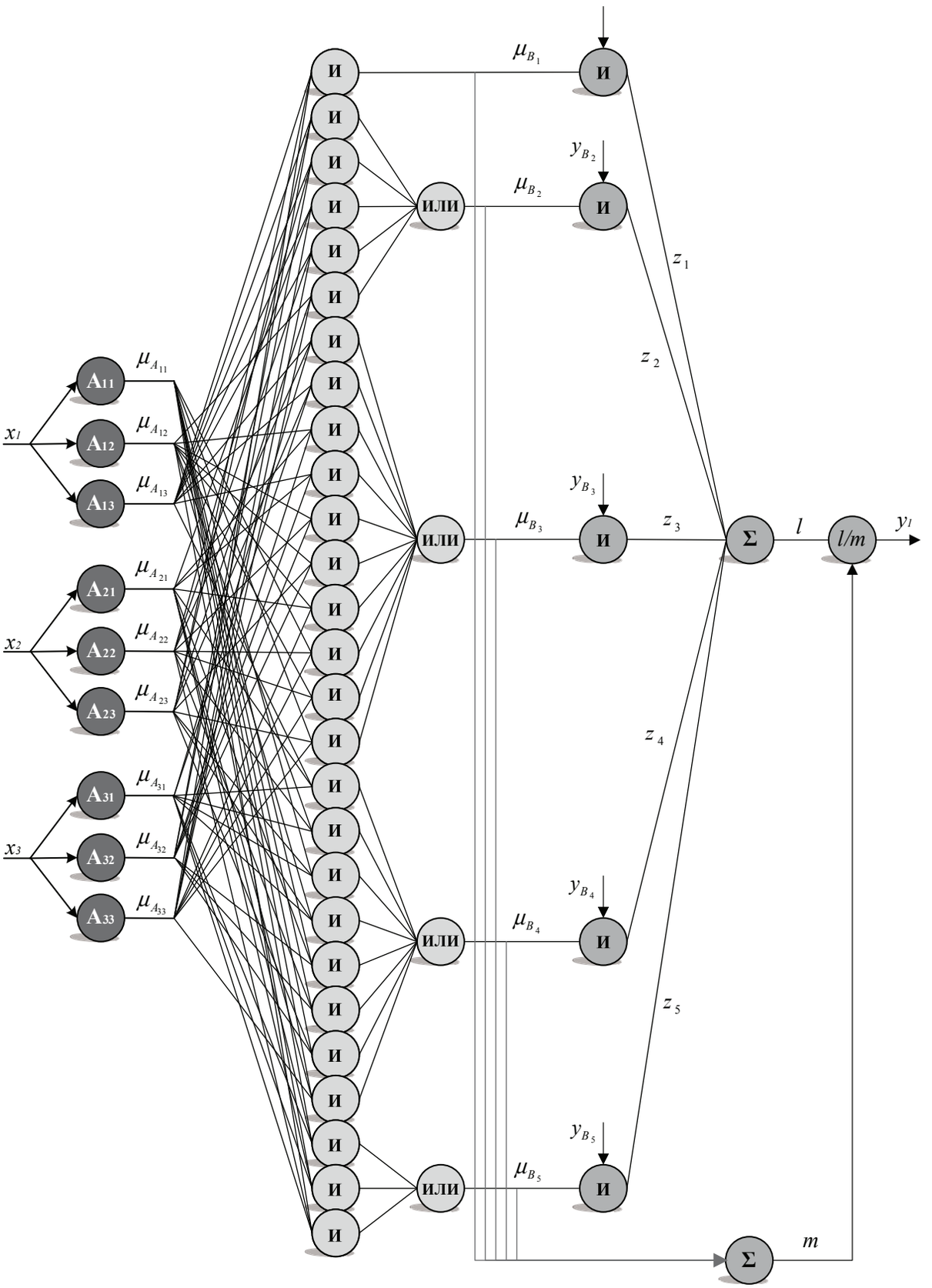

Рис. 3. Нейро-нечеткая сеть (фрагмент)

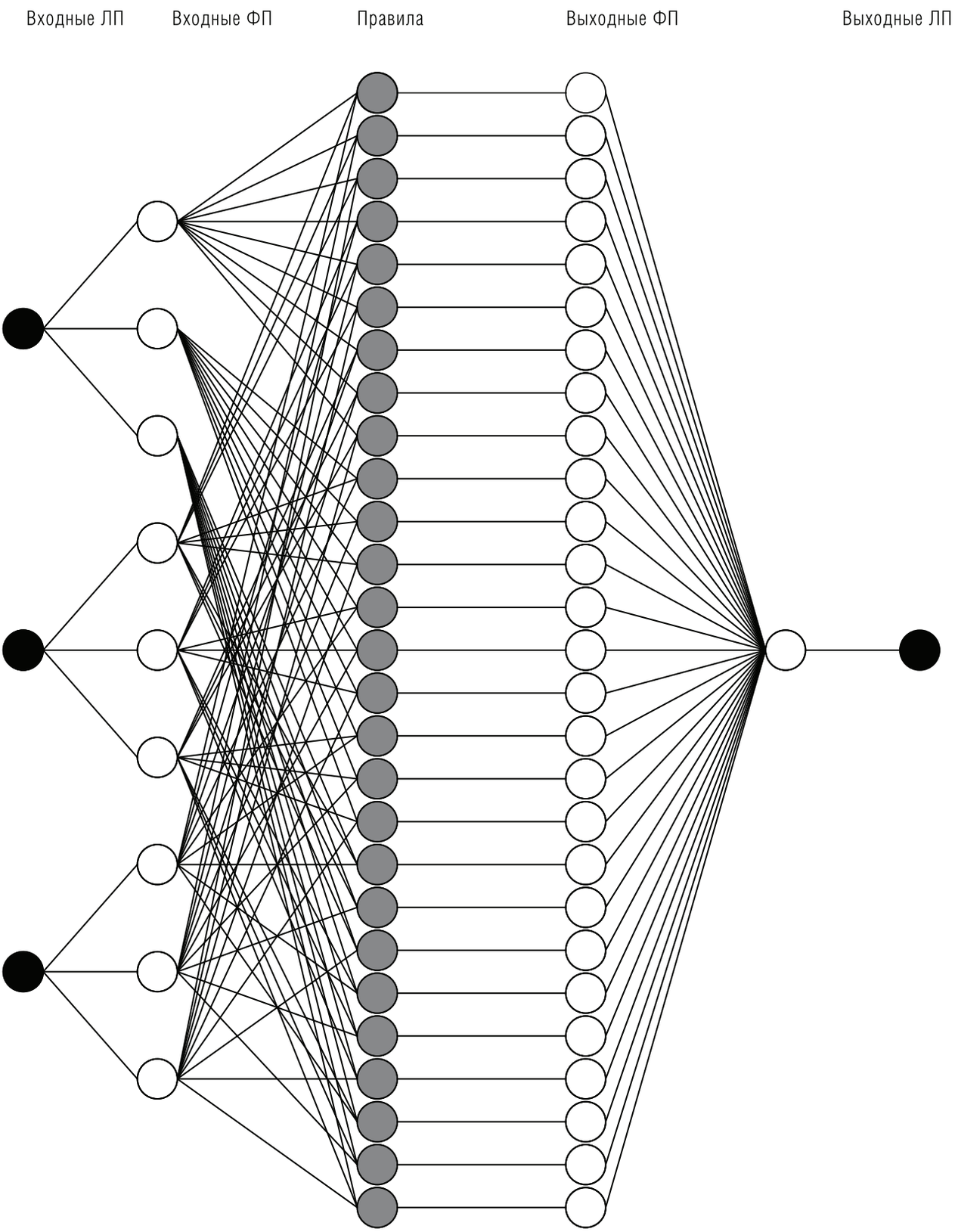

Рис. 5. Структура системы нечеткого вывода

тем применения метода Дельфи в рамках подхода, предложенного в работе [12]. Для формирования обучающих наборов также могут быть использованы данные, полученные из систем обнаружения вторжений, антивирусных программ, межсетевых экранов и других систем, включенных в СОИБ.

Пакет Neuro-Fuzzy Designer позволяет выполнять обучение методом обратного распространения ошибки, основным назначением которого является настройка всех слоев многослойной структуры путем изменения весов промежуточных слоев, и гибридным методом, который представляет собой комбинацию методов наименьших квадратов и обратного распространения ошибки. Результаты применения методов обучения ННС оценки рисков информационной безопасности приведены в *таблице 4.*

Таблица 4.

Применения методов обучения нейро-нечеткой сети

Метод обучения	Значение ошибки	Количество эпох[2]
Метод обратного распределения ошибки	0,0271	200
Гибридный метод	0,0108	28

Как видно из таблицы, гибридный метод обучения позволяет получить лучшие результаты значения ошибки сети за меньшее количество эпох. С учетом этого для настройки параметров функций принадлежности выбор был сделан в пользу гибридного метода. На *рисунке 6* приведен результат оптимизации (настройки) параметров функций принадлежности лингвистической переменной *ПрАнЗащ.*

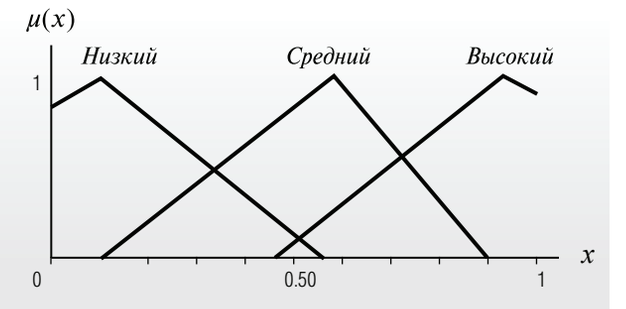

Рис. 6. Оптимизированные функции принадлежности

На *рисунке 7* приведена поверхность обученной нечеткой модели, которая показывает, как выходная ЛП зависит от двух входных ЛП, при этом значение третей переменной зафиксировано.

Графический вид зависимости выходной ЛП (*РискЗащ* – «Риск снижения эффективности защиты») от входных ЛП (*ПрАнЗащ* – «Программно-аппаратный уровень защиты» и *ОргЗащ* – «Уровень организационной защиты») показывает закономерный рост величины риска снижения эффективности защиты организации при уменьшении уровня программно-аппаратной защиты и уровня организационной защиты [2]. ■

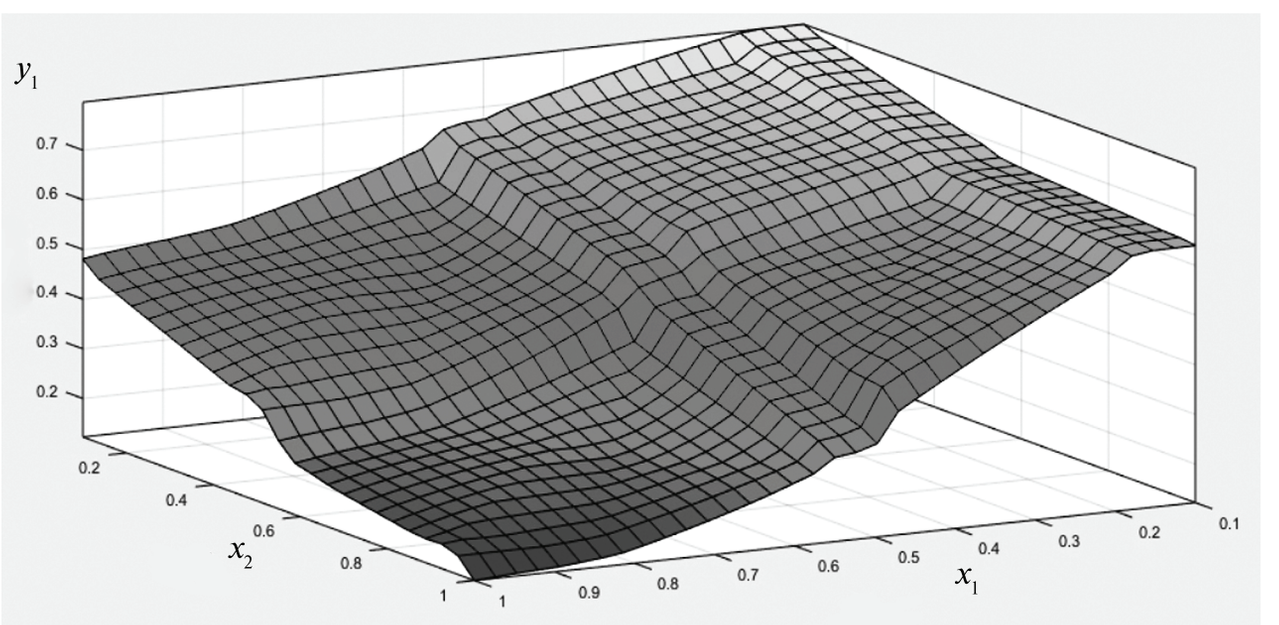

Рис. 7. Поверхность системы нечеткой модели

[2] Количество эпох, потребовавшееся для достижения указанного значения ошибки

Таким образом, гладкий и монотонный график зависимости приведенной «поверхности вывода» свидетельствует о хорошем «качестве» механизма вывода и о достаточности и непротиворечивости используемых правил вывода.

Механизм оценки рисков на основе ННС обладает широкими возможностями. В частности, он может быть адаптирован к имеющимся моделям управления рисками, а также модифицирован с учетом реальных условий политики информационной безопасности организации [7].

Заключение

Описанная во введении нечеткая продукционная модель являлась первым приближением для рассматриваемой предметной области и требовала настройки. Разработанная нечеткая самонастраивающаяся модель оценки рисков ИБ позволила провести корректировку параметров функций принадлежности лингвистических переменных для исследуемых систем обеспечения информационной безопасности и получить более адекватную нечеткую продукционную модель.

Реализованная нечеткая самонастраивающаяся модель позволяет проводить непрерывный анализа рисков ИБ, а полученные в результате нечеткого моделирования сведения позволяют ИТ-менеджерам определить приоритеты рисков (от «очень высокого» до «очень низкого») и разрабатывать эффективные планы мероприятий по снижению влияния наиболее опасных угроз. ■

Литература

1. Информационные технологии. Основные термины и определения в области технической защиты информации. Рекомендации по стандартизации (Р 50.1.053-2005). М., 2005.
2. Глушенко С.А. Применение системы MATLAB для оценки рисков информационной безопасности организации // Бизнес-информатика. 2013. № 4 (26). С. 35–42.
3. Risk management guide for information technology systems. Special publication 800-30. Gaithersburg, MD: NIST, 2002.
4. Симонов С.В. Анализ рисков, управление рисками // Jet Info. 1999. № 1 (68). С. 2–28.
5. Глушенко С.А. Нечеткая продукционная модель оценки рисков информационной безопасности организации // Вопросы экономики и права: Сборник статей аспирантов и соискателей ученой степени кандидата наук. Выпуск 11. Ростов-на-Дону: РГЭУ (РИНХ), 2013. С. 147.
6. Долженко А.И. Модель анализа риска потребительского качества проектов экономических информационных систем // Вестник Северо–Кавказского государственного технического университета. 2009. Т. 18. № 1. С. 129–134.
7. Глушенко С.А., Долженко А.И. Система поддержки принятия решений нечеткого моделирования рисков информационной безопасности организации // Информационные технологии. 2015. № 1. С. 68–74.
8. Борисов В.В., Круглов А.С., Федулов А.С. Нечеткие модели и сети. 2-е изд. М.: Горячая линия – Телеком, 2012. 284 с.
9. Пегат А. Нечеткое моделирование и управление. 2-е изд. М.: БИНОМ. Лаборатория знаний, 2013. 798 с.
10. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. М.: Горячая линия – Телеком, 2006. 452 с.
11. Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH. СПб: БХВ-Петербург, 2005. 736 с.
12. Хубаев Г.Н. Получение групповой экспертной оценки значений показателей: пошаговая процедура и программное обеспечение // Программные продукты и системы. 2011. № 2. С. 13–16.