УДК 004.056

Подход в использовании гибридных нейронных сетей при построении системы поддержки принятия решений для помощи в проведении аудита информационных систем персональных данных

Дмитриева Анастасия Витальевна
Амурский государственный университет
Студент

Аннотация
В приведённой статье рассматривается подход к построению модуля системы поддержки принятия решений, основанный на гибридной нейронной сети, проводящий оценку уровня защищённости исследуемой информационной системы персональных данных.
Ключевые слова: нечёткая логика, гибридная нейронная сеть, аудит, информационная система, система поддержки принятия решений

Approach in the use of hybrid neural networks in building a decision support system to assist in the audit for information systems of personal data

Dmitrieva Anastasia Vitalyevna
Amur State University
Student

Abstract
In this article, an approach is considered to build a module of a decision support system based on a hybrid neural network, assessing the level of security of the personal data information system under investigation
Keywords: fuzzy logic, hybrid neural network, audit, information system, decision support system

Информационные ресурсы в настоящее время являются одними из наиболее ценных ресурсов. Обеспечение их защиты всегда было довольно актуальной задачей. Понятие информационной безопасности сосредоточило в себе комплекс мероприятий, технических и программных средств, которые позволили так или иначе разрешить данных вопрос. Однако информационное общество развивается с каждым годом. Возникает необходимость в актуализации накопленных сведений, совершенствовании существующего законодательства, а также используемых технологий. Все информационные ресурсы, как правило, сосредоточены в специализированных системах – информационных системах, которые позволяют проводить обработку данных. В каждой такой системе должна быть обеспечена защита существующих данных в соответствии с нормативно-правовой документацией. Но не всегда нормативно-правовые акты содержат полную

базу, регламентирующие отношения в обработке персональных данных. Нередко приходится проводить анализ в условиях недостаточности исходных и промежуточных данных. Для таких ситуаций может послужить система поддержки принятия решений, которая поможет аудитору исследовать информационную систему персональных данных.

Процесс аудита является одним из механизмов обеспечения информационной безопасности. Аудит – процесс получения качественных и количественных оценок состояния информационной безопасности организации, компании, предприятия в соответствии с теми или иными критериями и показателями безопасности, рассмотренными в нормативно-правовой документации [1]. Аудит предназначен для оценки состояния информационной системы и выработки рекомендаций по применению комплекса организационных мер и программно-технических средств, направленных на обеспечение защиты информационных ресурсов информационной системы от существующих угроз информационной безопасности [2].

В соответствии с приведёнными выше определениями, можно выделить следующие цели проведения аудита:

1) анализ возможных рисков, связанных с возникновением угроз в отношении обработки содержащихся в информационной системе информационных ресурсов;

2) оценка текущего уровня защищённости информационной системы;

3) оценка соответствия текущего состояния защищённости требуемому в соответствии с нормативно-правовыми актами;

4) поиск и локализация уязвимостей в системе обеспечения безопасности;

5) выработка рекомендаций по усовершенствованию системы защиты.

Для одной из такой цели будет рассмотрен подход в использовании гибридных нейронных сетей. Сеть должна позволить провести оценку уровня защищённости исследуемой информационной системы персональных данных в условиях недостаточности данных предметной области.

Модуль системы поддержки принятия решений базируется на теории нечёткой логики. Нечёткая логика предназначена для формализации человеческих способностей к неточным или приближенным рассуждениям, которые позволяют более адекватно описывать ситуации с неопределённостью [3]. Математический раздел нечёткой логики в некотором роде обобщает разделы классической логики и теории множеств. Нечёткая логика основывается на понятии нечёткого множества, которое впервые было предложено Лютфи Заде в 1965 году.

Нечёткое множество – это множество, характеризующееся степенью неопределённости принадлежащих ему элементов, когда в тех или иных случаях невозможно с абсолютной уверенностью утверждать, что какой-либо элемент принадлежит данному множеству. Характеризует нечёткое множество функция принадлежности, которая говорит о степени принадлежности элемента к множеству. Функции принадлежности задаются

графически и имеют различные формы и виды. Например: треугольные, трапециевидные, Z-образные и S-образные, которые описываются своими математическими моделями. В теории нечёткой логики помимо понятия нечёткого множества используются понятия нечёткой переменной и лингвистической переменной. Лингвистическая переменная в свою очередь является обобщением нечёткой переменной, в которую помимо нечёткой переменной входит множество значений лингвистической переменной (термов лингвистической переменной), синтаксическая, а также семантическая процедура. Нечётким множествам могут быть свойственны аналогичные логические операции как в теории классической логики. К ним применимы такие операции, как например: конъюнкция, дизъюнкция, логическое отрицание, эквивалентность, импликация.

В теории нечёткой логики помимо всего прочего задействовано определение правила нечётких продукций. Правило нечётких продукций используется в системах искусственного интеллекта. Оно применяется для представления знаний и вывода определённых заключений предметной области в экспертных системах, а также для обработки, описания, анализа и моделирования сложных и слабо формализуемых процессов и систем [4]. Экспертные системы имеют дело с задачами искусственного интеллекта на более высоких уровнях, формируют управленческие решения с учётом сложившейся или прогнозируемой ситуации, накапливают эвристические знания и пытаются имитировать поведение эксперта [5].

Правило нечёткой продукции имеет следующий вид:

(i):Q;P;A=>B;S,F,N,

где (i) – наименование нечёткой продукции;

Q – область применения. Предназначается для описания исследуемой предметной области, которую выражает продукция;

P – условие применимости ядра нечёткой продукции. Логическое выражение (предикат), которое позволяет активизировать (обеспечить выполнение условий) ядро нечёткой продукции в случае истинности этого выражения;

A=>B – ядро нечёткой продукции, где:

A – условие ядра или посылка, или антецедент;

B – заключение ядра или консеквент;

«=>» – знак логического следования (или секвенции);

S – метод определения значения степени истинности заключения ядра. Данный метод в общем случае реализует алгоритм нечёткого вывода. Его также называют методом активизации или методом композиции;

F – коэффициент определённости нечёткой продукции. Также он называется весовым коэффициентом. Принимает значения в интервале [0, 1] и определяет количественную оценку степени истинности нечёткой продукции;

N – постусловия продукции. Описывает возможные действия при реализации ядра продукции.

Ядро продукции является центральным элементом данного понятия, который выражает послание в форме: «ЕСЛИ А, ТО В».

Совокупность нечётких продукций образует продукционную систему.

Исследуемый модуль системы поддержки принятия решений основан на процессе нечёткого вывода, который можно охарактеризовать как алгоритм получения нечётких заключений на основе нечётких условий (посылок). Он объединяет в себе понятия нечётких и лингвистических переменных, функций принадлежности и нечётких логических операций, таких, к примеру, как нечёткая импликация и нечёткая композиция. Системы нечёткого вывода являются частным случаем нечётких продукционных систем.

В общем случае алгоритм нечёткого вывода состоит из следующих этапов, схематически представленных на рисунке 1:

Рисунок 1 –Алгоритм нечёткого вывода

На этапе формирования правил создаётся множество правил нечётких продукций – база правил нечётких продукций. База правил должна быть полной, согласованной и непротиворечивой во избежание неадекватности полученного вывода.

На этапе фаззификации входных переменных вводят так называемую нечёткость. Здесь происходит процесс нахождения функций принадлежности лингвистических термов на основании обычных исходных данных.

На этапе агрегирования подусловий осуществляется определение степени истинности подусловий по всем правилам системы нечёткого вывода.

Этап активизации подзаключений реализует процесс нахождения степеней истинности подзаключений для каждого правила нечётких продукций.

На этапе аккумулирование заключений происходит нахождение функций принадлежности для каждой выходной лингвистической переменной. После чего следует дефаззификация выходных переменных или, иначе говоря, введение чёткости.

Данная теория использовалась при проектировании и реализации нейронных сетей. Нейронные сети представляют собой устройства параллельных вычислений, которые состоят из множества простых процессоров, где каждый из таких процессоров сети имеет дело только с сигналами, которые он в определённые временные промежутки получает, и сигналами, которые он с различной периодичностью посылает другим процессорам [6]. Основным их достоинством является то, что новую информацию о проблемной области можно получить на основе прогнозов. При этом они способны обучаться посредству уже имеющейся доступной информации.

Построение гибридной нейронной сети, основанной на теории нечёткого вывода, происходило в программной среде Matlab с использованием редактора Adaptive Neuro-Fuzzy Inference System (адаптивная система нейро-нечёткого вывода) – ANFIS.

Допустим, исследуемая информационная система персональных данных, которую необходимо проанализировать, классифицируется по 4-му уровню защищённости: категория обрабатываемых данных – общедоступные данные; количество субъектов персональных данных не превышает 100 000; третий тип актуальных угроз, связанный с отсутствием недекларированных возможностей в ОС и ПО. Информация о классификации информационной системы по тому или иному уровню защищённости содержится в Постановлении правительства Российской Федерации №1119. В этом же документе приведены требования, которые должны быть выполнимы в зависимости от уровня защищённости. К 4-му уровню защищённости приводятся 4 требования. Эти требования являются входными данными к гибридной нейронной сети:

1) режим обеспечения безопасности помещений, где обрабатываются ПДн (X1);

2) сохранность носителей (X2);

3) перечень лиц, допущенных к ПДн (X3);

4) средства защиты информации, прошедшие процедуру оценки соответствия требованиям законодательства РФ (X4) [7];

Необходимо на основании входных параметров определить лингвистические переменные и построить по ним систему правил, а также выборку для тестирования нейросети.

Лингвистические переменные X1, X2, X4 определяются неоднозначно. Для них приводится три лингвистических терма: S – низкий уровень выполнения требования; M – средний уровень выполнения требования; L –

высокий уровень выполнения требования. Для переменной X3, определяемой однозначно, приводится всего два терма: S – требование не выполняется; L – требование выполняется. Значения термов подбираются так, что для однозначно определяемой переменной они будут являться либо 0, либо 1 в зависимости от выполнения или невыполнения того или иного требования. Для неоднозначных переменных значения лингвистических термов выбираются из промежутков: для значения S – [0; 0,3), для значения M – [0,3; 0,9), для значения L – [0,9; 1]. Выходной параметр – оценка уровня защищённости определяется пятью лингвистическими термами: S – низкая оценка уровня защищённости; SM – оценка уровня защищённости ниже среднего; M – средняя оценка уровня защищённости; ML – оценка уровня защищённости выше среднего; L – высокая оценка уровня защищённости. Значения данных термов принадлежат промежуткам: S – [0; 0,1], SM – (0,1; 0,3], M – (0,3; 0,7), ML – [0,7; 0,9), M – [0,9; 1].

Фрагмент правил представлен на рисунке 2:

	A	B	C	D	E	F
1		Входные факторы				
2	№	X1	X2	X3	X4	Y
3	1	S	S	S	S	S
4						
5	2	S	S	L	M	M
6						
7	3	S	S	S	L	SM
8						
9	4	S	S	L	S	SM
10						
11	5	S	S	S	M	SM

Рисунок 2 – Фрагмент базы правил для нечёткой нейронной сети

Обучающая выборка создаётся на основе имеющихся сведений о лингвистических переменных. Здесь для каждой входной переменной случайным образом отбирается значение из приведённых интервалов, соответствующее тому или иному лингвистическому терму. Фрагмент обучающей выборки для тестирования нейросети изображён на рисунке 3:

№	X1	X2	X3	X4	Y
1	0.289	0.049	0	0.208	0.013
2	0.162	0.073	1	0.301	0.43
3	0.044	0.147	0	0.97	0.14
4	0.189	0.254	1	0.26	0.27
5	0.17	0.149	0	0.82	0.205
6	0.28	0.146	1	0.983	0.37
7	0.245	0.891	0	0.116	0.208
8	0.042	0.573	1	0.831	0.555
9	0.15	0.673	0	0.953	0.62
10	0.115	0.585	1	0.093	0.611

Рисунок 3 – Фрагмент обучающей выборки

Редактор ANFIS позволяет проектировать нейронную сеть. Он предусматривает различные инструменты для построения, генерирования и редактирования структуры сети. После проектирования структуры сети и её обучения необходимо провести тестирование.

В строке входных данных задаются значения параметров, для которых необходимо произвести оценку защищённости. Допустим, эксперт оценивает требования к уровням защищённости для конкретной ИСПДн следующим образом: X1 – 1 (L); X2 – 0,7 (M); X3 – 1 (L); X4 – 0,6 (M).

Тогда итоговый выходной показатель будет равен 0,8. Это значит, что данная информационная система персональных данных соответствует требованиям нормативно-правовой документации на 80%. Данная оценка говорит о том, что уровень защищённости системы выше среднего. Для того чтобы система выдавала наиболее высокий результат, следует увеличить значение одного из показателей (X2 или X4) с уровня M на уровень L. Результат работы нечёткой нейронной сети представлен на рисунке 4:

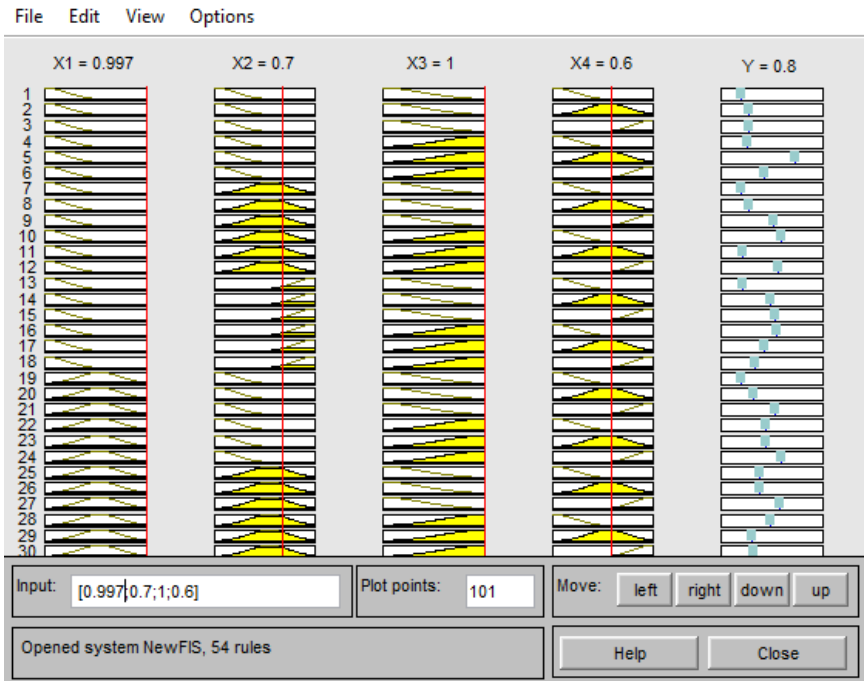

Рисунок 4 – результат работы нечёткой нейронной сети

Библиографический список

1. Петренко, С. А., Аудит безопасности Intranet: учебное пособие / С. А. Петренко, А. А. Петренко. М.:ДМК Пресс, 2002. 406 с.
2. Аверченков, В. И. Аудит информационной безопасности органов исполнительной власти [Электронный ресурс]: учебное пособие / В. И. Аверченков [и др.]. Электрон. текстовые данные. –Брянск: Брянский государственный технический университет, 2012. 100 с. Режим доступа: http://www.iprbookshop.ru/6992.html
3. Леоненков, А. В. Нечеткое моделирование в среде MATLAB и fuzzyTECH: учебное пособие / А. В. Леоненков. СПб, БХВ Петербург,

2005. 736 с.
4. Громов, Ю. Ю. Представление знаний в информационных системах: учебное пособие / Ю.Ю.Громов [и др.]. Тамбов: Тамбовский государственный технический университет, ЭБС АСВ, 2012. 169 с.
5. Аверченков, В. И. Основы математического моделирования технических систем [Электронный ресурс]: учебное пособие / В. И. Аверченков, В. П. Федоров, М. Л. Хейфец. Электрон. текстовые данные. Брянск: Брянский государственный технический университет, 2012. 271 с. Режим доступа: http://www.iprbookshop.ru/7003.html
6. Сысоев, Д. В. Введение в теорию искусственного интеллекта [Электронный ресурс]: учебное пособие / Д. В. Сысоев, О. В. Курипта, Д. К. Проскурин. Электрон. текстовые данные. Воронеж: Воронежский государственный архитектурно-строительный университет, ЭБС АСВ, 2014. 171 с. Режим доступа: http://www.iprbookshop.ru/30835.html
7. Постановление правительства «Об утверждении требований к защите персональных данных при их обработке в информационных системах персональных данных» №1119, утверждённое 1 ноября 2012 г: офиц. текст Москва: Кремль, 2012. – 4 с.