Вестник УГАТУ
Уфа : УГАТУ, 2011 Т. 15, № 5 (45). С. 54–65

УДК 004.891.2

В. И. Васильев, Н. В. Белков

СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПО ОБЕСПЕЧЕНИЮ БЕЗОПАСНОСТИ ПЕРСОНАЛЬНЫХ ДАННЫХ

Рассмотрены основные этапы создания системы поддержки принятия решений (СППР) по обеспечению безопасности персональных данных. Изложен подход к формированию базы знаний на основе правил и прецедентов с использованием онтологического анализа предметной области. Предлагается методика оценки эффективности принятых решений и анализа рисков информационной безопасности персональных данных. *Системы поддержки принятия решений; системы вывода по правилам*; *системы вывода по прецедентам*; *онтологический анализ*; *анализ рисков*

Федеральный закон «О персональных данных» был принят в Российской Федерации в июле 2006 года [1]. Появление данного закона, а также ряда подзаконных нормативных актов, вызвало резонанс в обществе. Проблема обеспечения безопасности персональных данных стала одной из самых обсуждаемых на различных профессиональных уровнях: управленческом, кадровом, юридическом и, конечно же, информационном. Об этом красноречиво свидетельствует количество проводимых в последние годы семинаров и конференций по данной тематике, а также всевозможных обучающих курсов для сотрудников организаций. Сложившаяся ситуация обусловлена особенностями обеспечения безопасности персональных данных в РФ.

АКТУАЛЬНОСТЬ ИССЛЕДОВАНИЯ

Область защиты персональных данных в Российской Федерации характеризуется рядом особенностей.

Во-первых, обширная сфера деятельности законодательства. Установленные требования обязательны для выполнения всеми операторами, т. е. организациями, осуществляющими обработку персональных данных. Учитывая, что ведение кадрового и бухгалтерского учета является неотъемлемой частью деятельности практически любого хозяйствующего субъекта, число операторов персональных данных достигает нескольких миллионов.

Во-вторых, нехватка квалифицированного персонала. Сравнительно небольшое число организаций всерьез подходит к проблеме обеспечения информационной безопасности, а значит, имеет в своем распоряжении соответствующие подразделения или специалистов. В то же время, только специалист в области информационной безопасности может разработать систему защиты, адекватную имеющимся угрозам безопасности.

В-третьих, несовершенство имеющейся нормативно-методической базы. Недостаточная проработанность, а зачастую и противоречивость положений принятых нормативных и методических документов затрудняет понимание требований и реализацию необходимых мероприятий. Кроме того, имеющиеся документы не затрагивают ряд важных мероприятий, таких как анализ рисков.

Таким образом, организации, приступающие к реализации мероприятий по защите персональных данных и не имеющие в своем штате специалистов по информационной безопасности, вынуждены либо направлять сотрудников на специализированные курсы, либо привлекать организации, оказывающие соответствующие услуги.

Курсы, как правило, длятся не более 5–7 дней, посвящены по большей части обзору законодательства и предоставляют лишь общие сведения о принципах и методах обеспечения информационной безопасности. Даже после обучения качественное выполнение всех необходимых мероприятий является для сотрудников трудновыполнимой задачей, требующей тщательного изучения предметной области. Привлечение сторонних организаций требует значительных материальных затрат, которые для крупных организаций исчисляются сотнями тысяч рублей.

Обеспечение безопасности персональных данных не разовое мероприятие, а непрерывный процесс, поэтому систему защиты необходимо постоянно поддерживать в актуальном состоянии. Среда, в которой осуществляется эксплуатация системы защиты персональных данных (СЗПДн), не является стационарной, поскольку со временем могут измениться:

Контактная информация: (347)273-06-72

- состав обрабатываемых персональных данных;
- организационная структура;
- бизнес-процессы;
- вычислительная сеть и др.

При внесении изменений в информационную систему необходимо проводить соответствующую корректировку параметров системы защиты, иначе ее эффективность значительно снижается или сводится на нет. Модификация системы защиты должна проводиться с привлечением специалистов.

Таким образом, в процессе формирования и эксплуатации системы защиты персональных данных необходимо принимать решения, требующие от лица, принимающего решения (ЛПР), знаний в области информационной безопасности, а также достоверных сведений об информационной системе, бизнес-процессах и информационных потоках в организации. Неверно принятое решение может привести к значительным финансовым потерям.

Вследствие этого, одной из ключевых задач является поддержка принятия решений для обеспечения достаточного уровня защищенности персональных данных.

МЕТОДИКА РАЗРАБОТКИ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Общая схема этапов разработки системы поддержки принятия решений (СППР) представлена на рис. 1.

Этап *идентификации задач* СППР является предварительным, подготовительным этапом разработки системы. На данном этапе определяются основные задачи, которые должна решать СППР, ее функции и общая структура. Также на данном этапе осуществляется сбор и накопление документации, которая содержит информацию о предметной области, используемую в процессе принятия решений [2].

На этапе *моделирования предметной области* осуществляется построение комплекса системных моделей с использованием структурного подхода.

На этапе *формирования информационного пространства представления знаний* проводится онтологический и семантический анализ процессов создания системы защиты персональных данных.

Разрабатываемая СППР является интеллектуальной информационной системой, поэтому одним из основных этапов является этап *разработки базы знаний*. На данном этапе проводится определение прецедентов и правил поддержки принятия решений, а также формирование и подготовка необходимого набора данных и знаний [2].

Рис. 1. Этапы разработки СППР

Реализация системы поддержки принятия решений подразумевает написание программного кода, а также физическое наполнение базы знаний. На этом же этапе осуществляется реализация прототипов системы.

На заключительном этапе производится *оценка эффективности* построенной СППР.

ИДЕНТИФИКАЦИЯ ЗАДАЧ СППР ПРИ ОРГАНИЗАЦИИ ЗАЩИТЫ ПЕРСОНАЛЬНЫХ ДАННЫХ

В результате анализа процессов обеспечения безопасности персональных данных были определены основные задачи системы поддержки принятия решений. Можно выделить три основных задачи, решаемые СППР:

- накопление и систематизация сведений об информационной системе, в которой осуществляется обработка персональных данных;
- поддержка принятия решений (ППР) при проектировании системы защиты персональных данных;
- поддержка принятия решений при эксплуатации системы защиты персональных данных.

Каждая из выделенных задач, в свою очередь, может быть разложена на несколько подзадач. Так, в рамках задачи поддержки принятия

решений при проектировании необходимо решение следующих подзадач: ППР при выделении и классификации информационных систем персональных данных (ИСПДн); ППР при определении угроз безопасности персональных данных и степени их актуальности; ППР при определении необходимых средств защиты; проведение анализа рисков.

Общая структура СППР, отображающая наиболее общие принципы ее работы, представлена на рис. 2.

Работа с СППР осуществляется следующим образом. ЛПР вводит в СППР сведения, касающиеся изменений в информационной системе, и формирует запрос на получение рекомендаций либо проведение анализа рисков. На основе полученных сведений СППР определяет, каким образом изменения сказались на структуре ИСПДн, угрозах безопасности персональных данных (УБПДн) либо используемых средствах защиты, и формирует соответствующие рекомендации. Если предложенные рекомендации утверждаются ЛПР, то изменения вносятся в базу данных.

Для формирования рекомендаций предлагается совместное использование механизмов вывода на основе правил и на основе прецедентов: если решение не удается найти с помощью имеющихся правил, то инициируется поиск схожей ситуации в базе прецедентов. Если же поиск по базе прецедентов не дал результата, то выдается соответствующее сообщение пользователю с предложением сформировать новый прецедент. Таким образом происходит обучение базы знаний.

МОДЕЛИРОВАНИЕ ПРЕДМЕТНОЙ ОБЛАСТИ

Системное моделирование является основой процесса создания интеллектуальных информационных систем [3]. Для построения системных моделей использовалась технология структурного анализа и проектирования (SADT) и, в частности, технология IDEF. С ее помощью был построен комплекс функциональных, информационных и динамических моделей.

На рис. 3 представлена функциональная модель IDEF0 на одном из уровней иерархии (рис. 3, *а*), а также фрагмент информационной модели IDEF1X (рис. 3, *б*).

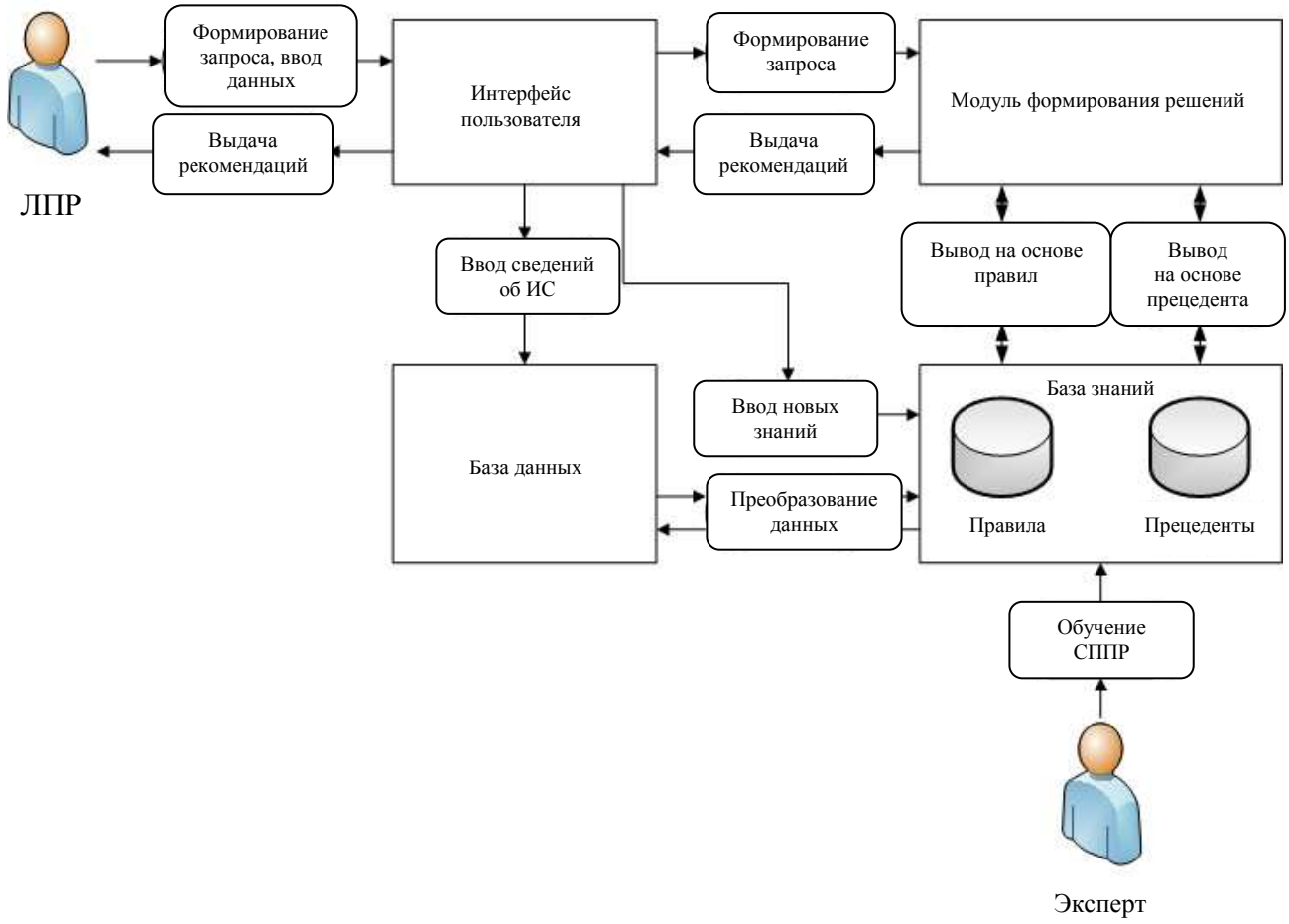
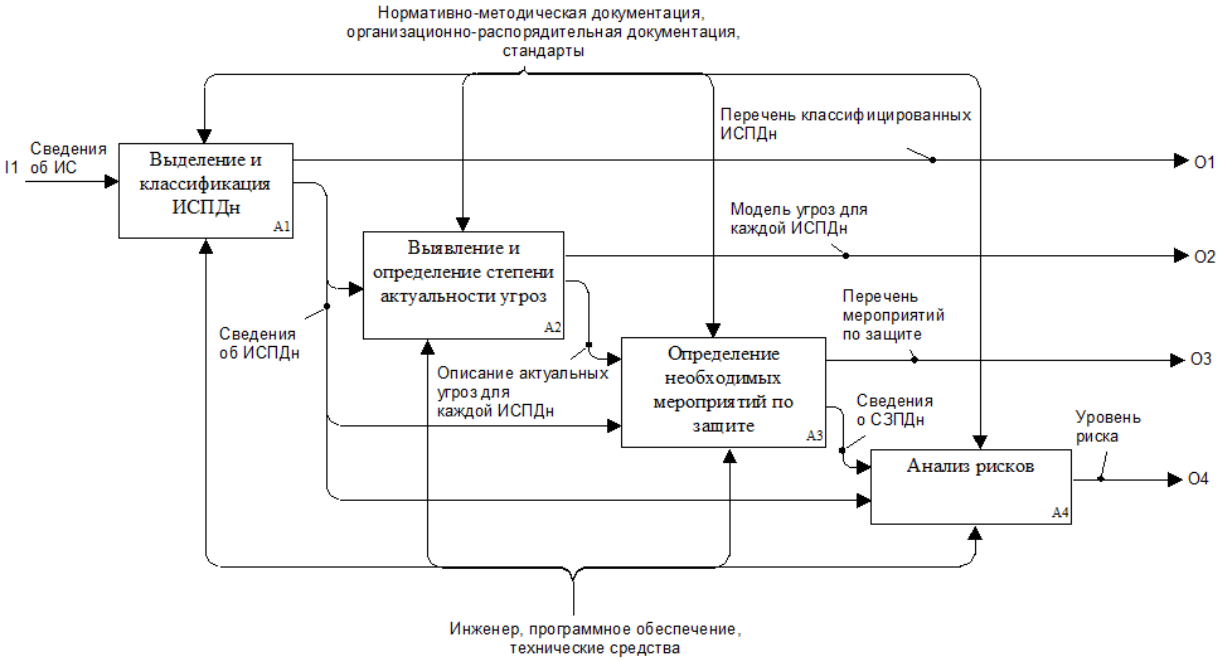
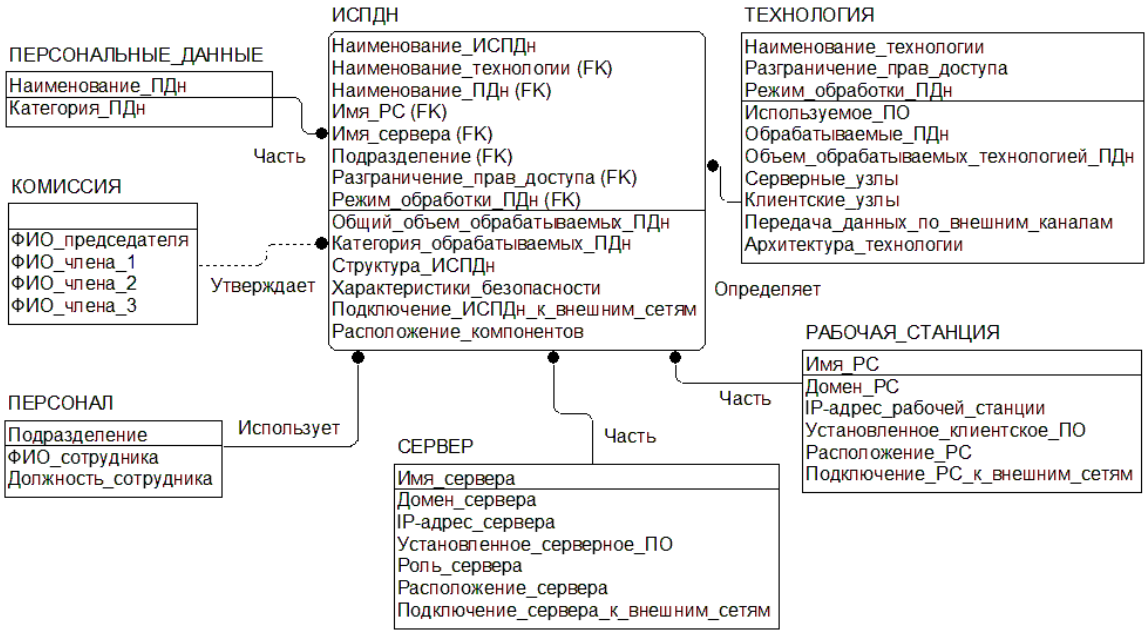


Рис. 2. Общая структура СППР



а



б

Рис.3. Примеры структурных моделей IDEF:
а – функциональная модель IDEF0; *б* – информационная модель IDEF1X

ФОРМИРОВАНИЕ ИНФОРМАЦИОННОГО ПРОСТРАНСТВА ПРЕДСТАВЛЕНИЯ ЗНАНИЙ

В слабоформализованных областях, таких как область защиты персональных данных, недостаточно комплекса системных моделей, так как они не позволяют достаточно полно отразить семантику процессов управления и принятия решений, а также выделить существенные понятия и связи между ними. В таких случаях принято использовать онтологический подход [4]. Онтология служит для обеспечения общего словаря для человека и системы поддержки принятия решений, позволяя, таким образом, им взаимодействовать. Наиболее распространенное определение онтологии, которое дал Том Грубер [5], звучит следующим образом: «*Онтология* – формальная спецификация разделяемой концептуализации». Концептуализация подразумевает описание множества объектов (понятий), знаний о них и связей между ними.

Компонентами онтологии являются:

- множество понятий предметной области и их атрибутов;
- множество отношений, или ассоциаций между выделенными понятиями;
- множество аксиом и правил вывода, заданных на выделенных понятиях.

Таким образом, модель онтологии [4, 5] можно представить как упорядоченную тройку вида

$$O = <T, R, F>, \quad (1)$$

где T – конечное множество концептов (понятий, терминов) предметной области; R – конечное множество отношений между концептами; F – конечное множество аксиом, заданных на концептах.

При этом, если множества R и F пустые, т. е. $R = \emptyset$ и $F = \emptyset$, то онтология O трансформируется в *словарь* (V):

$$V = <T>.$$

Если же пустое только множество F, т. е. $R \neq \emptyset$, F = $\emptyset$, то онтология является *тезаурусом* (Th), состоящим из множества концептов и множества отношений.

$$Th = <T, R>.$$

В случае, если множество R включает в себя единственный тип отношений «быть элементом класса», то тезаурус становится *таксономией*, отображающей иерархию понятий.

Приведенная модель онтологии показывает последовательность действий при формировании онтологии. Сначала разрабатывается словарь терминов, затем на полученных терминах (концептах) определяются отношения иерархии, эквивалентности, исключения, формируя таким образом тезаурус. В заключение определяются функции интерпретации (аксиомы).

При разработке словаря терминов было принято решение использовать три источника знаний: системные модели, полученные в результате моделирования; документация области обработки и защиты персональных данных (нормативно-методическая, организационно-распорядительная, проектно-техническая документация); а также знания экспертов в данной предметной области.

Словарь системных моделей содержит наименования сущностей и их атрибутов.

Для выделения терминов из нормативно-методической и проектно-технической документации производился лингвистический анализ текста при помощи лингвистического процессора «Text Analyst 2.0». Средствами данного программного продукта анализировался каждый исходный документ. В результате были составлены словари терминов. Фрагмент словаря, построенного по результатам анализа методического документа «Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных» [6], представлен в табл. 1. Для каждого термина словаря указывается частота появления термина в анализируемом документе, его вес, показывающий, насколько важную роль играет понятие для смысла всего текста, а также иные термины, наиболее тесно связанные с исходным.

Составленные словари были предоставлены экспертам предметной области для анализа. В результате этого были удалены малозначащие термины, а также добавлены новые, не выделенные из документации в ходе лингвистического анализа и в процессе моделирования. Итоговый словарь является основной составляющей тезауруса и онтологии предметной области. Дальнейшее формирование тезауруса и онтологии осуществляется в программной оболочке «Protégé» на языке OWL DL (Ontology Web Language on Description Logic).

Каждый термин из словаря вносится в онтологию в качестве отдельного класса, таким образом, словарь переносится в «Protege». Затем, путем перемещения классов, формируется иерархия понятий, т. е. таксономия, представленная на рис. 4. Понятие, расположенное по дереву классов ниже, связано с родительским понятием отношением наследования «быть элементом класса».

Таблица 1

Термин	Частота	Вес	Связанные термины
ИСПДн	259	100	Сеть, ПДн, нарушитель, сеть международного информационного обмена, технические средства, сети связи общего пользования, УБПДн, безопасность
ПДн	120	100	ИСПДн, безопасность, обработка, угроза, УБПДн
Сеть	228	100	ИСПДн, сеть международного информационного обмена, подключение, сеть связи общего пользования, ПДн, нарушитель, хост
Автоматизированное рабочее место	17	99	ПДн, подключение, сеть, сеть международного информационного обмена, сеть связи общего пользования

На следующем шаге между понятиями определяются отношения эквивалентности и исключения. Отношение эквивалентности (синонимии) задается в разделе «Equivalent classes» путем указания всех эквивалентных классов. Так, например, для понятия «Информационная система персональных данных» эквивалентным является понятие «ИСПДн», а для понятия «Идентификатор» – понятие «Код». Отношение исключения задается в разделе «Disjoint classes».

Завершает построение онтологии задание множества аксиом. Аксиомы ассоциируют идентификаторы классов с частичной или полной спецификацией их характеристик, предоставляют информацию о классах, на основании которых можно определить непротиворечивость и осуществить вывод на онтологии [5]. Часть аксиом содержится в самой таксономии, описывающей отношения обобщения. Именно аксиомы отображают правила, сформированные в онтологии [4].

ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ ПРАВИЛ

В разрабатываемой базе знаний используются две модели представления знаний: в форме правил и в форме прецедентов.

Правила – это способ представления основных знаний предметной области, которые объясняют возникновение тех или иных явлений, дают возможность прогнозировать развитие ситуации, позволяют связывать отдельные объекты реального мира [4]. Правила отображают модель рассуждения эксперта в ситуациях, по которым накоплено достаточно примеров принятия решений. В виде правил представляются знания, логическая система которых упорядочена.

С целью поддержки принятия решений в области обеспечения безопасности персональных данных выделяются следующие категории правил:

- правила принятия решений:
- правила выделения ИСПДн;
- правила классификации ИСПДн;
- правила определения угроз безопасности;
- правила определения актуальности угроз;
- правила выбора мероприятий по защите ПДн.
- правила построения базы знаний;
- правила адаптации прецедентов;
- правила оценки эффективности и анализа рисков.

Наиболее простые правила имеют вид:

$$R = <A_1, A_2, \ldots, A_n; B>, \qquad (2)$$

где $A_1 \ldots A_n$ – предпосылки срабатывания правила; B – заключение.

Такое представление правила соответствует продукционной модели. Продукционное правило есть выражение вида: «**ЕСЛИ** A, **ТО** B». Левая часть правила (антецедент A) содержит набор предпосылок, правая часть (консеквент B) – заключение правила. Правило срабатывает, если выполняются все предпосылки. Каждая предпосылка A_i представляет собой простое выражение «атрибут-значение», построенное с использованием терминов из онтологии предметной области. Между собой предпосылки связываются при помощи элементарных операций И, ИЛИ, НЕ.

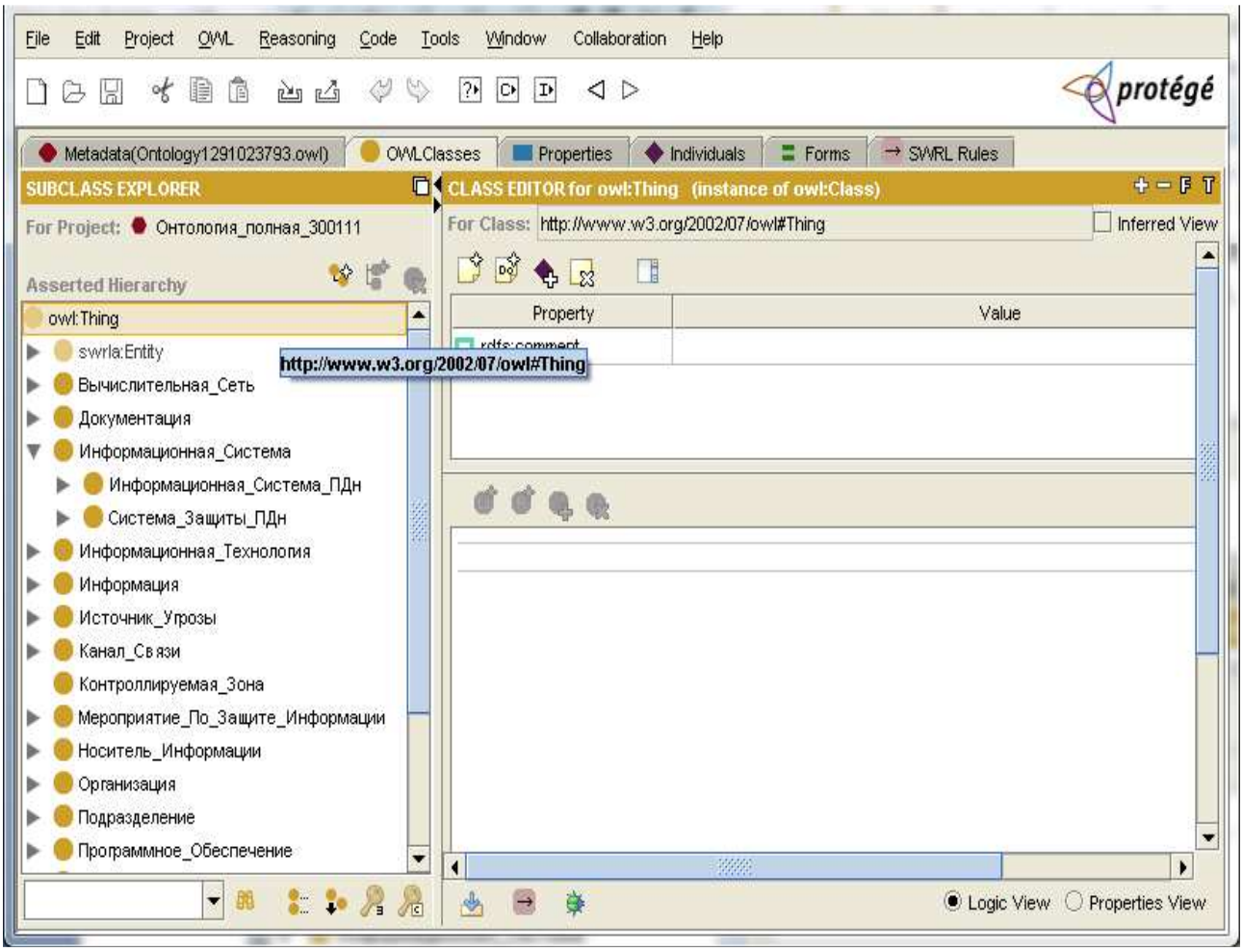


Рис. 4. Рабочее окно программы «Protégé» со сформированной таксономией

Ниже приведены примеры правил классификации ИСПДн.

R1: Если Число_серверов(ИСПДн) = 0 И Число_РС(ИСПДн) = 1, то Структура(ИСПДн) = 'Автономная';

R2: Если ((Число_серверов(ИСПДн) ≥ 1) ИЛИ (Число_РС(ИСПДн) > 1)) И (Передача_по_общим_сетям(ИСПДн) = false), то Структура(ИСПДн) = 'Локальная';

R3: Если ((Число_серверов(ИСПДн) ≥ 1) ИЛИ (Число_РС(ИСПДн) > 1)) И (Передача_по_общим_сетям(ИСПДн) = true), то Структура(ИСПДн) = 'Распределенная'.

Представленные правила применимы в условиях полноты и определенности имеющихся знаний, т. е. когда содержащиеся в них утверждения абсолютно достоверны. Тем не менее, процессу принятия некоторых решений по обеспечению безопасности персональных данных присуща неопределенность.

В случае задачи определения угроз безопасности неопределенность вносится путем введения в правила степеней уверенности. При этом формулировка правила (2) усложняется и принимает следующий общий вид:

$$R = <A_1, U_1, A_2, U_2, \ldots, A_n, U_n; B, U_B>,$$

где U_i – степень уверенности в предпосылке A_i; U_B – степень уверенности в заключении.

Для расчета степеней уверенности предлагается использовать меру доверия Демпстера–Шефера [7].

При решении задачи определения актуальности угроз наиболее уместным является применение подхода на основе нечеткой логики. Это связано с тем, что при определении степени исходной защищенности, частоты (вероятности) реализации и опасности угроз достаточно сложно определить точные числовые значения. Сформированные нечеткие правила представляют собой импликацию вида:

Если x_1 есть A_1 **И** x_2 есть A_2 **И** … x_n есть A_n, **ТО** y есть B,

где $A_1 \ldots A_n$, B – лингвистические переменные; $x_1 \ldots x_n$ – входные переменные; y – выходная переменная.

Пример нечеткого правила:

Если Возможность_реализации_угрозы есть «низкая» **И** Опасность_угрозы есть «средняя», **ТО** Актуальность_угрозы есть «неактуальная».

Для реализации детерминированных правил предлагается использовать языки логического программирования, например, Prolog либо Semantic Web Rule Language (SWRL). Формирование нечетких правил осуществляется с помощью нейро-нечетких сетей в среде MATLAB с использованием соответствующего модуля.

ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ ПРЕЦЕДЕНТОВ

В ситуациях, когда использование базы правил не позволяет получить решение, либо правила еще не сформулированы ввиду отсутствия достаточных примеров принятия решении, применяется база прецедентов [2]. При использовании вывода на основе прецедентов используются знания о предыдущих ситуациях или случаях (прецедентах).

Формальная модель системы вывода, основанного на прецедентах, представлена в виде упорядоченной тройки [2]:

$$CBR = <Case, Onto, Retr>,$$

где Case – база прецедентов; Onto – онтология предметной области; Retr – алгоритм поиска прецедентов.

Для организации эффективного поиска в базе важно провести индексацию и классификацию прецедентов. В рассматриваемой СППР база прецедентов используется в задачах выбора мероприятий по защите персональных данных, а также выбора дальнейших действий при изменении информационной системы. По этой причине множество прецедентов Case классифицировано в соответствии с классами подсистем защиты персональных данных (подсистемы разграничения доступа, антивирусной защиты и др.) и классами проблемных ситуаций (изменение организационной структуры, изменение состава обрабатываемых ПДн и др.).

Последовательность основных процедур вывода на основе прецедентов приведена на рис. 5.

Первоначально формируется новый прецедент, описывающий сложившуюся проблемную ситуацию. Создание прецедента осуществляется на основе имеющихся в базе данных сведений об информационной системе в целом, а также об ИСПДн. В случае необходимости сведения уточняются у пользователя. Прецедент должен содержать описание проблемы, принятые для решения проблемы действия и результаты применения решения. Соответственно, прецедент можно представить в виде совокупности следующих объектов [2]:

$$Case_k = <Case\_name_k, C_i, X_k^i, D_k, E_k>, \quad (3)$$

где $Case\_name_k$ – идентификатор прецедента; C_i – класс, к которому относится прецедент; X_k^i – множество значений признаков, составляющих описание проблемной ситуации и образующих входной вектор; D_k – множество возможных решений, содержащихся в прецеденте; E_k – множество оценок эффективности принятых решений.

Прецеденты представлены в базе знаний в виде отдельных фреймов. Новый прецедент содержит только слоты $Case\_name_k$, C_i, X_k^i.

Рис. 5. Основные процедуры вывода, основанного на прецедентах

После того, как новый прецедент создан, начинается поиск ближайших прецедентов в данном классе, основанный на аналогии. При этом происходит сравнение вектора признаков текущего прецедента и прецедента из базы.

Для сравнения прецедентов был выбран метод «ближайшего соседа». Недостатком данного метода является низкое быстродействие, однако эксплуатация СППР в режиме дефицита времени не планируется. Кроме того, поиск осуществляется не по всей базе, а только внутри определенного класса, что сокращает затраченное время. При использовании метода «ближайшего соседа» для каждого признака X_k^i из X^i вычисляется мера сходства, в качестве которой используется взвешенная евклидова метрика:

$$\mathrm{sim}(X_k, X_j) = \sqrt{\sum_{i=1}^{N} \omega_i (X_k^i - X_j^i)^2}, \quad (4)$$

где X_k^i и X_j^i – значение i-го признака для k-го и j-го прецедента соответственно; N – общее

количество параметров для данного класса прецедентов; ω_i – вес i-го признака.

В случае, если признак является качественным либо логическим, его значение нормализуется. Веса признаков определяются экспертами с использованием метода анализа иерархий. Ближайшим является прецедент, метрика которого окажется минимальной.

После того, как ближайший прецедент найден, его решение адаптируется к проблемной ситуации при помощи соответствующих правил. Полученный в результате прецедент после оценки его эффективности заносится в базу знаний. В случае же, если ближайший прецедент найти не удается, то описание проблемной ситуации сохраняется как нерешенный инцидент. Пользователю выдается соответствующее сообщение с предложением найти решение самостоятельно, после чего новый прецедент заносится в базу. Таким образом происходит обучение системы.

ОЦЕНКА ЭФФЕКТИВНОСТИ ПРИНЯТЫХ РЕШЕНИЙ. АНАЛИЗ РИСКОВ

Для проведения оценки эффективности принятых решений и анализа рисков предложена следующая методика.

Наиболее распространенная оценка информационных рисков имеет вид:

$$R = p \cdot C,$$

где R – информационный риск; p – вероятность нарушения информационной безопасности; C – стоимость информационных ресурсов (активов).

При построении системы защиты персональных данных достаточно сложно рассчитать общее значение вероятности нарушения безопасности, но возможно найти значения вероятности реализации отдельных угроз. Соответственно, общее значение риска можно найти как сумму значений рисков от реализации всех угроз:

$$R = \sum_{i=1}^{n} R_i = \sum_{i=1}^{n} p_i \cdot L_i, \qquad (5)$$

где R_i – информационный риск от реализации i-й угрозы; p_i – вероятность реализации i-й угрозы; L_i – ущерб от реализации i-й угрозы.

Каждый класс угроз в системе поддержки принятия решений по обеспечению безопасности персональных данных может быть представлен в виде совокупности следующих объектов:

$$T = <T\_\text{name}, S, V, W, O, A>, \qquad (6)$$

где T_name – название класса угроз; S – множество возможных источников угроз данного класса; V – множество возможных уязвимостей, связанных с данным классом угроз; W – множество возможных способов реализации угроз данного класса; O – множество возможных объектов воздействия угроз данного класса; A – множество возможных деструктивных действий, оказываемых на ИСПДн при реализации угроз данного класса.

Содержание перечисленных множеств для каждого класса угроз определяется экспертами и содержится в базе знаний.

Множества S, V, O, A для угрозы отдельной ИСПДн составляются путем выбора из множества допустимых для данного класса угроз значений, определенных экспертами, в соответствии со свойствами и условиями функционирования ИСПДн. Множество W формируется в зависимости от состава других множеств. При этом множество A состоит из трех подмножеств:

$$A = <A_к, A_ц, A_д>,$$

где $A_к$ – множество деструктивных действий, направленных на нарушение конфиденциальности ПДн; $A_ц$ – множество деструктивных действий, направленных на нарушение целостности ПДн; $A_д$ – множество деструктивных действий, направленных на нарушение доступности ПДн.

Большинство известных методик и программных продуктов анализа рисков предполагают введение значения вероятности реализации угрозы p_i экспертом, т. е. пользователем системы, на основе каких-либо статистических сведений. В случае анализа угроз безопасности персональных данных подобная статистика на предприятии, как правило, отсутствует. Кроме того, использование статистических сведений не позволяет определить, как изменяется значение при внедрении того или иного механизма защиты. По этой причине предлагается метод вычисления вероятностей на основе сведений, содержащихся в описании угрозы (6).

Ключевым при вычислении вероятности реализации угрозы является множество возможных способов ее реализации W. Чем больше возможных способов реализации угрозы, тем выше вероятность того, что злоумышленник попытается реализовать данную угрозу. Как уже было сказано, множество W формируется в соответствии с тем, какие возможны источники угрозы, какие уязвимости и объекты присутствуют в ИСПДн. Таким образом, вероятность реализации i-й угрозы безопасности для конкретной ИСПДн может быть найдена по формуле (7).

$$p_i = \frac{\sum_{k=1}^{m} \alpha_k \cdot pr_k}{m}, \qquad (7)$$

где pr_k – вероятность успешного использования злоумышленником k-го варианта реализации i-й угрозы; α_k – коэффициент, определяющий степень уверенности в том, что злоумышленник может воспользоваться k-м вариантом реализации i-й угрозы; m – количество возможных способов реализации i-й угрозы.

Коэффициенты α_k определяются ЛПР на основе имеющейся статистической информации. Если статистическая информация отсутствует, то все варианты реализации угрозы считаются равновероятными; коэффициент α_k принимает два возможных значения: $\alpha_k = 0$, если отсутствуют объективные предпосылки для использования рассматриваемого варианта реализации угрозы, и $\alpha_k = 1$, если таковые предпосылки имеются.

Каждый вариант реализации угрозы может быть полностью перекрыт определенным набором мероприятий по защите. При этом вероятность успешного использования варианта реализации угрозы зависит от того, какое количество возможных мероприятий реализовано, и может быть рассчитана по следующей формуле:

$$pr_k = 1 - \frac{mr_k}{mt_k}, \qquad (8)$$

где mr_k – число реализованных мероприятий, перекрывающих k-й вариант реализации угрозы; где mt_k – общее число мероприятий, перекрывающих k-й вариант реализации угрозы.

Расчет размера ущерба от реализации i-й угрозы безопасности j-й ИСПДн производится по следующей формуле:

$$L_j^i = (\omega_{\text{кон}}^j + \omega_{\text{цел}}^j + \omega_{\text{дост}}^j) \cdot C_j, \qquad (9)$$

где $\omega_{\text{кон}}^j$, $\omega_{\text{цел}}^j$, $\omega_{\text{дост}}^j$ – весовые коэффициенты, определяющие насколько критичным является обеспечение конфиденциальности, целостности и доступности в j-й ИСПДн; C_j – стоимость информационных ресурсов, содержащихся в j-й ИСПДн.

Значения весовых коэффициентов задаются экспертами для каждого типа ИСПДн и могут быть изменены ЛПР.

Ниже представлен пример расчета значения риска для угрозы «Анализ сетевого трафика».

Для данной угрозы в общем случае определены шесть возможных вариантов реализации Р1–Р6 ($m = 6$):

Р1: Использование вредоносной программы;

Р2: Внедрение программно-аппаратной закладки;

Р3: Перехват информации, передаваемой по внутренним каналам связи, с использованием уязвимостей протоколов канального уровня;

Р4: Перехват информации, передаваемой по внутренним каналам связи, с использованием уязвимостей протоколов сетевого уровня;

Р5: Перехват информации, передаваемой по внешним каналам связи, с использованием уязвимостей протоколов сетевого уровня;

Р6: Использование уязвимостей прикладного и специального программного обеспечения.

Пусть в рассматриваемой ИСПДн вычислительная сеть построена с использованием коммутаторов, без использования сетей беспроводного доступа и без передачи данных по внешним каналам связи. Статистическая информация отсутствует. В этом случае объективные предпосылки для использования вариантов Р3 и Р5 отсутствуют, а коэффициенты имеют значения:

$$\alpha_3 = \alpha_5 = 0;\ \alpha_1 = \alpha_2 = \alpha_4 = \alpha_6 = 1.$$

Для перекрытия оставшихся вариантов реализации угрозы (Р1, Р2, Р4, Р6) выделены следующие мероприятия:

Для Р1: Средства антивирусной защиты.

Для Р2: Средства обеспечения целостности и физическая охрана внутренних каналов связи.

Для Р4: Средства обнаружения вторжений.

Для Р6: Средства обеспечения целостности и средства анализа защищенности.

Если в ИСПДн не реализованы перечисленные мероприятия, то вероятность успешного использования всех вариантов реализации угрозы равна 1. В этом случае вероятность реализации угрозы определяется по формуле (7):

$$p = \frac{1+1+1+1}{6} = \frac{4}{6} = 0{,}6667.$$

При применении в ИСПДн средств обеспечения целостности, вероятности успешного использования вариантов реализации угрозы принимают значения:

$$pr_1 = pr_4 = 1;\ pr_2 = pr_6 = 1 - \frac{1}{2} = 0{,}5.$$

Тогда вероятность реализации угрозы принимает значение:

$$p = \frac{1+0{,}5+1+0{,}5}{6} = \frac{3}{6} = 0{,}5.$$

Таким образом, имеется возможность рассчитать значение вероятности реализации угрозы при использовании различных средств защиты и их сочетаний.

Для рассматриваемой ИСПДн экспертами были определены следующие значения коэффициентов, определяющих чувствительность обрабатываемых в ней персональных данных к нарушениям характеристик безопасности:

$$\omega^i_{\text{кон}} = 0{,}7;\ \ \omega^i_{\text{цел}} = 0{,}2;\ \ \omega^i_{\text{дост}} = 0{,}1\,.$$

Реализация угрозы «Анализ сетевого трафика» может привести лишь к нарушению конфиденциальности информации, поэтому размер ущерба от реализации данной угрозы принимает значение:

$$L^i = \omega^i_{\text{кон}} \cdot C = 0{,}7 \cdot C\,,$$

где C – стоимость защищаемых информационных ресурсов.

В табл. 2 приведены примеры расчета вероятностей реализации угрозы «Анализ сетевого трафика» и относительная величина ущерба от ее реализации при применении различных мероприятий по защите персональных данных и некоторых их сочетаний.

Таблица 2

Реализованные мероприятия	Вероятность реализации угрозы	Уровень риска
Отсутствуют	0,667	0,467C
Средства антивирусной защиты	0,5	0,35C
Средства обеспечения целостности	0,5	0,35C
Физическая охрана внутренних каналов связи	0,583	0,408C
Средства обнаружения вторжений	0,5	0,35C
Средства анализа защищенности	0,583	0,408C
Средства антивирусной защиты + средства обеспечения целостности	0,333	0,233C
Средства антивирусной защиты + физическая охрана внутренних каналов связи + средства обнаружения вторжений	0,25	0,175C

Представленная методика анализа рисков позволяет определить эффективность принятых или планируемых решений, а также оценить эффективность прецедентов. Кроме того, значение вероятности реализации угрозы зависит от числа имеющихся на ее пути барьеров, что соответствует определенному методикой ФСТЭК порядку.

Методика на данный момент предполагает ряд допущений, например, не учитывает эффективность перекрытия тем или иным мероприятием какого-либо способа реализации угрозы. Кроме того, необходимо дополнительно описать методику определения стоимости информационных ресурсов, содержащих персональные данные. Тем не менее данная методика позволяет оценить эффективность и корректность генерируемых решений по результатам тестирования прототипа СППР.

Модуль анализа рисков и оценки эффективности, наряду с модулями вывода на основе прецедентов и вывода на основе правил, является основным компонентом разрабатываемой системы поддержки принятия решений.

Использование данной СППР позволит организациям и предприятиям проводить классификацию ИСПДн, формировать модель угроз, проектировать систему защиты персональных данных, а также поддерживать ее в актуальном состоянии, не имея в своем штате специалистов в области информационной безопасности. Результатом применения СППР является повышение эффективности принимаемых решений и снижение затрат при создании и эксплуатации системы защиты персональных данных.

ВЫВОДЫ

Дается описание системы поддержки принятия решений по обеспечению безопасности персональных данных. Показаны актуальность проблемы и наличие потребности в подобной системе.

Определены требования, предъявляемые к СППР, ее структура, а также методика проектирования и создания. Выделены ключевые задачи, решаемые системой.

Предложена и подробно описана методика формирования информационного пространства представления знаний предметной области обеспечения безопасности персональных данных на основе онтологического анализа.

Предложено построение базы знаний с использованием модулей вывода на основе правил и на основе прецедентов. Подобная структура базы знаний позволяет использовать как формализованные знания, представленные в виде правил, так и знания, основанные на предыдущем опыте и отображаемые в виде прецедентов.

Предложена методика оценки эффективности принятых решений и анализа рисков информационной безопасности ИСПДн, позво-

ляющая с помощью разработанной СППР проводить исследования в условиях отсутствия статистической информации по угрозам безопасности.

СПИСОК ЛИТЕРАТУРЫ

1. Федеральный закон Российской Федерации от 27 июля 2006 г. № 152-ФЗ "О персональных данных" (с изменениями от 27.12.2009 № 363-ФЗ, от 28.06.2010 № 123-ФЗ, от 27.07.2010 № 204-ФЗ).

2. Поддержка принятия решений при стратегическом управлении предприятием на основе инженерий знаний / Л. Р. Черняховская [и др.]. Уфа: АН РБ, Гилем, 2010. 128 с.

3. Автоматизированное проектирование информационно-управляющих систем. Проектирование экспертных систем на основе системного моделирования / Г. Г. Куликов [и др.]. Уфа, УГАТУ, 1999. 223 с.

4. **Бадамшин Р. А., Ильясов Б. Г., Черняховская Л. Р.** Проблемы управления сложными динамическими объектами в критических ситуациях на основе знаний. М.: Машиностроение, 2003. 240 с.

5. **Башмаков И. А., Башмаков А. И.** Интеллектуальные информационные технологии. М.: Изд-во МГТУ им. Н. Э. Баумана, 2005. 304 с.

6. Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных: Метод. документ Федеральной службы по техническому и экспортному контролю (ФСТЭК России).

7. **Муромцев Д. И.** Введение в технологию экспертных систем. СПб.: СПб ГУ ИТМО, 2005. 93 с.

ОБ АВТОРАХ

Васильев Владимир Иванович, проф., зав. каф. вычислит. техники и защиты информации. Дипл. инженер по промэлектронике (УГАТУ, 1970). Д-р техн. наук по системн. анализу и авт. упр. (ЦИАМ, 1990). Иссл. в обл. многосвязн., многофункц. и интеллектуальн. систем.

Белков Николай Вячеславович, асп. той же каф. Дипл. специалист по защите информации (УГАТУ, 2009). Готовит дис. в обл. защиты информации и поддержки принятия решений.