ЛОГИКО-ЛИНГВИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ УГРОЗ БЕЗОПАСНОСТИ ИНФОРМАЦИИ В ИНФОРМАЦИОННЫХ СИСТЕМАХ

Язов Ю.К[1]., Соловьев С.В.[2],Тарелкин М.А.[3]

Цель статьи: *оценка возможностей применения и сравнительная характеристика реляционных языков логико-лингвистического моделирования для формализованного описания процессов реализации угроз безопасности информации в информационных системах.*

Метод: *применение аппарата логико-лингвистического моделирования, позволяющего формально описать угрозы безопасности информации и совокупность действий, выполняемых в ходе их реализации, с учетом возможностей реляционных языков описания, таких как язык Кодда, контекстно-свободный плекс-язык, язык RX-кодов, синтагматических цепей и семантических сетей.*

Полученный результат: *дана краткая характеристика и сравнительный анализа реляционных языков логико-лингвистического моделирования, указаны особенности, влияющие на возможности их применения для описания угроз безопасности информации и процессов их реализации. Показана целесообразность такого моделирования при создании перспективных экспертных систем (систем поддержки принятия решения при организации и ведении технической защиты информации в информационных системах), предназначенных для автоматизированного и автоматического анализа угроз, при ведении банка данных угроз по результатам мониторинга публикаций о них в сети Internet.*

Приведены примеры построения формальных логико-лингвистических описаний широко известных угроз проведения компьютерных атак на информационные системы с использованием языков RX-кодов, синтагматических цепей и семантических сетей, даны предложения по построению на основе реляционных языков специализированного языка логико-лингвистического описания угроз, в том числе в условиях применения мер и средств защиты от них.

Отмечено, что предложенный подход к логико-лингвистическому моделированию угроз безопасности информации применим в случае отсутствия необходимости учета фактора времени при оценке возможностей их реализации, однако позволяет формировать исходные данные для такой оценки.

Ключевые слова: *реляционный язык, оценка, функциональная модель, сеть Петри-Маркова, мера защиты.*

DOI:10.21681/2311-3456-2022-4-13-25

1. Введение

Огромное разнообразие угроз безопасности информации в информационных системах (ИС) и способов их реализации обусловили необходимость создания электронных баз данных для сбора и накопления сведений об угрозах, их актуализации и применения для решения задач защиты информации в ИС. Сегодня эти сведения представляют собой вербальные описания, что практически исключает их использование в автоматизированном режиме в формальных процедурах и алгоритмах анализа угроз. Вместе с тем уже сегодня крайне востребованы для автоматизации процессов прогнозирования и анализа угроз соответствующие экспертные системы [1 – 3], основанные на использовании элементов теории искусственного

1 Язов Юрий Константинович, доктор технических наук, профессор, главный научный сотрудник управления ФАУ «ГНИИИ ПТЗИ ФСТЭК России», г. Воронеж, Российская Федерация. E-mail: Yazoff_1946@mail.ru

2 Соловьев Сергей Вениаминович, начальник управления ФАУ «ГНИИИ ПТЗИ ФСТЭК России», г. Воронеж, Российская Федерация. E-mail: gniii@fstec.ru

3 Тарелкин Михаил Андреевич, старший научный сотрудник ФАУ «ГНИИИ ПТЗИ ФСТЭК России», г. Воронеж, Российская Федерация. E-mail: gniii@fstec.ru

интеллекта, таких как искусственные нейронные сети, нечеткие множества и нечеткая логика, генетические алгоритмы, эволюционная семиотика [4] и др. Переход от вербальных к формальным описаниям угроз безопасности информации связан с необходимостью разработки соответствующих правил формального описания. Весьма перспективным для этого является использование логико-лингвистических моделей [5]. До сих пор логико-лингвистические описания использовались преимущественно для решения задач поиска разнородной информации [6], построения онтологических баз знаний [7] и семантического анализа контента в сетях общего пользования [8, 9], а также при решении отдельных частных вопросов обеспечения безопасности информации, таких как оценивание защищенности на основе семантической модели метрик и данных [10], автоматизация анализа уязвимостей программного обеспечения [11], семантический анализ информационных рисков и угроз [12 – 15] и др.

Если в вербальных моделях применяются естественные языки описания, то в логико-лингвистических моделях – символьные описания с использованием специально созданных искусственных языков.

В соответствии с [5] в классификационную схему искусственных языков, как правило, включают три группы языков:

1) логические языки, которые используются в формальных системах логического типа, таких как исчисление высказываний или исчисление предикатов первого порядка, многозначные логики и модальные исчисления;

2) реляционные языки, для которых характерно введение конечного множества бинарных отношений, с помощью которых передаются бинарные связи между элементами языка;

3) языки ролевых фреймов, в которых каждая единица в предложении описывается через семантические роли, которые она может выполнять.

При использовании этих языков могут применяться формализмы математической логики, теории нечетких множеств, теории графов и др. Строгость логических отношений между элементами описания может варьироваться в широких пределах: от отношений, выраженных нечеткими оценками, до отношений строгого детерминизма. Наиболее часто упоминаются в литературе и используются на практике так называемые реляционные языки (от англ. relation – отношение) и языки ролевых фреймов [5].

Следует отметить, что, хотя фреймовые языки представления знаний предоставляют пользователю определенную свободу при описании знаний, так как допускают различные способы описания данных в пределах одного фрейма, их недостатками являются, во-первых, возможность применения для решения сравнительно несложных задач описания, так как при расширении описания фреймовая сеть имеет свойство разрастаться до значительных размеров и проблемы поиска решения в таких сетях становятся трудноразрешимыми, поскольку связи между фреймами в сетях, описывающих объемные знания, как правило, неоднозначны. Во-вторых, фреймовые сети менее приспособлены к адаптации, так как внесение новых фреймов и изменение слотов (элементов описания) в имеющихся фреймах может повлечь противоречия и зацикливания в ссылках при движении по иерархической структуре фреймовой сети. В связи с этим возможности языков фреймов для описания угроз безопасности информации в ИС далее не рассматриваются.

В данной статье излагается подход к использованию реляционных языков для формального описания угроз безопасности информации в ИС.

2. Краткая характеристика реляционных языков для логико-лингвистического моделирования угроз

К реляционным языкам, как правило, относят табличные языки Кодда, контекстно-свободный плекс-язык, тензорный язык Крона[4], RX-коды, синтагматические цепи и семантические сети, [5]. Краткая характеристика этих языков приведена в таблице (табл.1), кроме тензорного языка, который оказывается весьма сложным в практическом применении при описании угроз.

Анализ табличных языков Кодда и, контекстно-свободного плекс-языка показал, что их весьма сложно использовать при логико-лингвистическом описании угроз безопасности или из-за недостаточных выразительных возможностей, а тензорный язык Крона оказывается весьма сложным в использовании даже при описании весьма простых угроз.

Рассматривая широко известные *RX-коды*, необходимо отметить, что в этом языке в качестве правильных синтаксических выражений используются двойки из множества понятий $\{X_j^m\}$ и множества отношений $\{R_i\}$ (см. табл.1) и конъюнкции двоек. Для интерпретации выражений языка RX-кодов всем понятиям и отношениям предписываются некоторые семантические значения, например, для понятий: X_i^1 – «угроза

4 Сегодня известно достаточно много и других языков, однако их возможности для описания угроз безопасности информации весьма ограничены и далее не рассматриваются.

$$X_0^0 = (R_1 X_1^1)(R_1 X_2^1)(R_1 X_3^1)(R_1 X_5^1) \,\&\, \left\{ R_2 X_4^1 R_3 \left[R_1 \left(X_5^1 \right) \right] \hat{\vee} \; R_3 X_6^1 R_4 X_5^2 R_1 X_7^1 \right\}$$

Имеют место (R_1) класс угроз - «Отказ в обслуживании» (X_1^1); имя угрозы -переполнение буфера (X_2^1); Источник угрозы - внешний нарушитель (X_3^1); Состояние буфера операционной системы до атаки нормальное (X_5^1)	Повторяется (R_2) отправка хакером пактов с запросом на соединение (X_4^1), пока выполняется условие (R_3), при котором имеет место (R_1) нормальное состояние буфера операционной системы (X_5^1)	При наличии условия (R_3), когда количество отправленных хакером пакетов превышает предельное для буфера атакуемого компьютера значение (X_6^1) следствием (R_4) становится состояние переполнение буфера (X_5^2) и имеет место (R_1) отказ в обслуживании (X_7^1)

Рис.1. Пример формальной записи угрозы сетевой атаки «syn-flood» с использованием RX-кодов

перехвата трафика с использованием «сниффера», относящаяся с номером 1 к i-му классу угроз, объединяющему все угрозы перехвата трафика; X_j^1 – «угроза переполнения буфера операционной системы», относящаяся с номером 1 к j-му классу угроз «Отказ в обслуживании».

Семантические значения для отношений могут быть, например, следующие: R_m – «иметь имя»; R_k - «повторять» с порядковым номером k; R_n - «быть меньше» с порядковым номером n и т.д. Приведем пример использования языка RX-кодов для описания угрозы атаки «sin-flood».

Пример 1. Суть атаки «syn-flood» заключается в посылке большого количества пакетов сообщений, сформированных по протоколу TCP с запросом на соединение с атакуемым компьютером (TCP-запросов). Атакуемый компьютер по технологии установления соединений направляет на атакующий компьютер хакера квитанцию о получении запроса и согласии на установление виртуального канала, при этом записывает данные запросившего компьютера в буфер. Атакующий, получив эту квитанцию, должен был бы направить свое уведомление о том, что он начинает передачу, которая на атакуемой стороне является сигналом для очищения буфера и начала приема пакетов. Однако хакер не посылает это уведомление, а направляет следующий запрос, что заканчивается очередной записью данных об атакующем компьютере в буфер атакуемого. Это многократно повторяется, что приводит к переполнению буфера атакуемого компьютера и его «зависанию».

Пусть рассматриваемая угроза обозначается как X_0^0 и в описании угрозы используются следующие понятия: X_1^1 - имеет место класс угроз «Отказ в обслуживании»; X_2^1 – имя угрозы «Переполнение буфера операционной системы»; X_3^1 – источником угрозы является внешний нарушитель; X_4^1 – пакет с запросом на соединение отправлен; X_5^1 – состояние буфера операционной системы атакуемого компьютера нормальное; X_5^2 – состояние буфера операционной системы атакуемого компьютера соответствует отказу в обслуживании; X_6^1 – количество отправленных хакером пакетов а запросом на соединение превышает предельное для буфера атакуемого компьютера значение; X_7^1 – состояние компьютера «отказ в обслуживании».

Суть отношений между понятиями состоит в следующем: R_1 – имеет место; R_2 – выполняется повтор; R_3 – выполняется условие; R_4 – следствием является.

Для расширения возможностей описания могут использоваться пропозициональные связки, такие как операция конъюнкции $\&$ (соответствующая логике «И»), операция дезъюнкции в разделительном смысле $\hat{\vee}$ (соответствующая логике «ИЛИ», когда утверждается одно и только одно из высказываний), и др.

Простейшая формальная запись угрозы атаки «syn-flood» с использованием RX-кодов показана на рисунке (рис.1).

Язык *синтагматических цепей* (СЦ) предоставляет значительно больше возможностей для описания угроз безопасности информации, чем язык RX-кодов. Это обусловлено не только тем, что построение синтагм интуитивно понятно, но и тем, что для расширения

его возможностей используются квантификаторы, модификаторы и императивы[5]. Еще более важным

5 Императивы – словесные выражения указаний, реализуемых в ИС в интересах, например, парирования угроз или решения других задач защиты. Примерами императивов являются такие слова и словосочетания, как «провести идентификацию», «изменить пароль доступа» и др. Квантификаторы служат для введения количественных и качественных мер для понятий, отношений и императивов. Примерами идентификаторов могут быть слова «всегда», «для всех», «очень редко» и т.п. Модификаторы приписывают понятиям, отношениям и императивам некоторые характеристики, которые в них ранее не фиксировались. Примерами квантификаторов могут быть слова «опасная», «аварийное», «зависание» и т.п.

Таблица 1

Краткая характеристика реляционных языков описания [2]

№	Наименование языка	Форма записи информации на языке	Достоинства и недостатки
1	Табличный язык Кодда	Данные о предметной области в виде строки таблицы формально задаются на множестве: $H = I \times H_1 \times ... \times H_k$, где H_1 – множества значений атрибутов j_1, а I – множество имен сущностей, данных в виде n-арных отношений ($n \le k$)	На заре информатизации язык Кодда с табличной формой представления информации составлял теоретическую основу почти всех методов разработки всевозможных баз данных. С точки зрения записи и обработки знаний этот язык представляет сегодня скорее исторический интерес, но на практике он используется в качестве базиса для построения различных баз данных и развития языков более высокого уровня. Язык позволяет описывать данные, но не знания, поэтому его применение для логико-лингвистического описания угроз нецелесообразно[1]
2	Контекстно-свободный плекс-язык	Основными компонентами этого языка являются: плекс-элементы (абстрактный объект, имеющий определенное количество контактов, то есть входов и выходов для соединения с другими плекс-элементами), операция конкатенации и грамматика, определяющая правила соединения плекс-элементов между собой. Плекс-элементы образуют алфавит языка. Операция конкатенации формально задается матрицей, элементы которой отражают связи между контактами i и j, то есть матрицей $A\|\gamma_{ij}\|$. Грамматикой языка называется четверка: $G = \langle v_0, A, V, R \rangle$, где v_0 – начальный символ, $A = \{a_i^p\}$ – алфавит терминальных плекс-элементов; $V = \{v_k\}$ – алфавит вспомогательных плекс-элементов; R – множество правил вывода [2]	Основным достоинством языка является возможность его использования как достаточно гибкой формы представления структур различных объектов в базах знаний. Причем, когда появляется некоторая новая структура, всегда можно установить, достаточно ли алфавита и грамматики данного языка для ее адекватного представления в базе знаний, или необходимо дополнять алфавит новыми плекс-элементами и модифицировать грамматику языка. На наш взгляд, применение этого языка для описания угроз безопасности информации нецелесообразно
3	Язык RX-кодов	Язык относится к так называемым дескрипторным языкам*. Базовыми единицами этого языка выступают отношения и понятия. Отношения, называемые парадигмастическими, т.е. отношениями, существующими всегда, обозначаются буквой R_i с различными нижними индексами. Понятия обозначаются буквой X_j^m с верхними и нижними индексами. В качестве правильных синтаксических выражений используются двойки R_i и X_j^m, любые конъюнкции таких двоек и конъюнкции, получающиеся путем замены любого X в правильном синтаксическом выражении. Знак конъюнкции, если это не приводит к путанице, обычно опускается, а для указания зоны действия отношения используются круглые скобки. Эти правила и есть синтаксис языка RX-кодов. Для интерпретации выражений языка RX-кодов всем понятиям и отношениям жестко предписываются некоторые семантиче-	Естественная иерархия, присущая RX-кодам, позволяет описывать знания о предметной области методом древовидных структур, однако одновременно накладывает существенные ограничения на их описательные возможности. Дело в том, что определяемое в этом языке понятие всегда является корнем дерева, а отношения между понятиями одного уровня не учитываются. Для сочленения одноуровневых понятий требуется введение дополнительных (внеязыковых) средств. Вместе с тем именно это ограничение позволяет весьма эффективно использовать язык RX-кодов при организации баз знаний, в которых родовидовые отношения играют определяющую роль. Для естественного языка древовидные конструкции не столь характерны. В нем отношения между понятиями имеют, как правило, сетевую структуру и содержат многочис-

* В дескрипторных языках отражаются только понятия и имена, которые называют дескрипторами.

№	Наименование языка	Форма записи информации на языке	Достоинства и недостатки
3	Язык RX-кодов	ские значения. В этом случае оказывается возможным производить вывод новых понятий через ранее определенные, которые для этого нового понятия выступают в качестве признаков	ленные циклы, в которых понятия определяются через самих себя. Поэтому при использовании RX-кодов для записи естественно-языковых текстов возникают определенные трудности. Фактически, RX-коды и построенные на их основе методы представления знаний не могут адекватно отражать тексты, написанные на естественном языке. К основным недостаткам RX-кодов относят следующие: — определяемое в этом языке понятие всегда является корнем дерева, а отношения между понятиями одного уровня не учитываются, при этом для сочленения одноуровневых понятий требуется введение дополнительных (внеязыковых) средств, что затрудняет использование языка; — имеется определенная сложность написания запросов и медленные и сложные алгоритмы поиска записей в базах знаний; — может иметь место ложная координация дескрипторов (понятий и отношений). С учетом изложенного язык RX-кодов находит ограниченное применение на практике. Вместе с тем он может быть использования для описания несложных угроз безопасности информации в соответствующих базах знаний о них
4	Язык синтагматических цепей	Язык *синтагматических цепей* является, по сути, расширением языка RX-кодов. В основе синтагматических цепей лежат элементарные синтагмы, представляющие собой тройки вида xry, где x – понятия (понятие-класс, понятие-процесс или понятие-состояние), y – имена, r – отношение между ними. Например, если в средней позиции находится отношение «иметь имя», то в левой позиции стоит понятие, для которого в правой позиции указывается имя	В отличие от RX-кодов синтагматическая цепь состоит из «троек», определяющих отношение между понятиями и именами. Это существенно облегчает построение формальной записи. К недостаткам относятся сложность, громоздкость описания совокупности фактов, учет зависимости понятий, сложность вывода новых (производных) понятий. Вместе с тем синтагматические цепи так же, как и RX-коды, могут быть использованы для описания несложных угроз безопасности информации.
5	Язык семантических сетей	Формально семантические сети задаются тремя классами термов: понятиями $X=\{x_1,x_2,\ldots,x_n\}$, именами $I=\{i_1,i_2,\ldots,i_m\}$ и отношениями $R=\{\rho,r_1,r_2,\ldots,r_k\}$. Отношение ρ – особое. Оно означает «иметь имя», остальные отношения подразделяются на падежные, характеристические, причинно-следственные, иерархические, временные и топологические. *Падежные отношения* связывают предикат как основу действия, определяемого предложением, с остальными словами. *Характеристические отношения* связывают характеристику с характеризуемым объектом и характеристику со значением характеристики. *Причинно-следственные отношения* связывают понятия, одно из которых является причиной, а другое	При соответствующем выборе обозначений с помощью семантических сетей можно выразить очень сложные совокупности фактов. В отличие от RX-кодов и синтагматических сетей семантические сети позволяют описывать не только постоянные отношения, присущие совокупности объектов, но и временные (ситуативные) отношения между объектами. Важная особенность языка семантических сетей заключается в его способности выводить новые понятия (обобщать ситуации) за счет выделения типовой структуры отношений. Эта особенность не только используется для описания ситуаций – она оказалась весьма продуктивной при ситуационном поиске решений. Постепенное обобщение описаний в языке

DOI:10.21681/2311-3456-2022-4-13-25

№	Наименование языка	Форма записи информации на языке	Достоинства и недостатки
5	Язык семантических сетей	следствием. *Иерархические отношения* указывают на то, что один объект является составной частью другого объекта или – одно понятие определяется через другие понятия. *Временные отношения* бывают двух типов: абсолютные и относительные. Абсолютные временные отношения устанавливаются между объектами и отрезками (точками) временной оси, а относительные – это связки «быть раньше», «одновременно», «быть позже» и др. *Топологические отношения* связывают объекты с точками какой-либо системы координат либо указывают на их взаимное расположение: «быть впереди», «быть сзади», «располагаться левее» и так далее	семантических сетей приводит, в конце концов, к построению описания с максимальной свободой в заполнении перемен-ных конкретными понятиями и со свободными параметрами. Это важное свойство семантических сетей позволяет использовать их при создании интеллектуальных баз знаний. К недостаткам семантических сетей можно отнести следующие: во-первых, они представляют собой пассивные структуры, для обработки которых необходим специальный аппарат формального вывода; во-вторых, представление, использование и модификация знаний при описании сложных угроз остается достаточно трудоемкой процедурой, особенно при наличии множественных отношений между ее элементами описания. Однако эти недостатки вполне преодолимы.

является расширение состава элементов описания процесса реализации угроз такими классами функциональности как модальность[6], что можно использовать в алгоритмах функционирования перспективных систем поддержки принятия решений по ТЗИ. Рассмотрим описание той же атаки «syn-flood» с использованием языка синтагматических цепей.

Пример 2. Введем следующие понятия и отношения: A_0 – атака «syn-flood»; b_1 – компьютер хакера; b_2 – атакуемый компьютер; b_3 – запрос на соединение; b_4 – буфер атакуемого компьютера; b_5 – квитанция; b_6 – уведомление; n_1 – количество передаваемых на атакуемый компьютер пакетов с запросами на соединение; n_2 – предельно допустимое количество записей в буфер атакуемого компьютера о запросившем соединение компьютере; s_1 – состояние буфера нормальное; s_2 – переполнение буфера (зависание); r_1 – имеет имя; r_2 – передает, r_3 – блокирует; r_4 – получает; r_5 – записывает; r_6 – находится в состоянии; r_7 – повторяет; r_8 – больше; r_9 – стало причиной; i_0 – класс атаки; i_1 – TCP-пакет с запросом на соединение; i_2 – TCP-пакет с квитанцией о согласии на соединение; i_3 – сведения о компьютере, запросившем соединение, записываемые в буфер; k_0 – «Отказ в обслуживании». Тогда формальная запись данной атаки имеет следующий вид:

$$A_0 = (i_0 r_1 k_0) \cdot \left\{ b_2 r_7 \left[(b_2 r_6 s_1) \left((b_1 r_2 i_1)(b_2 r_4 i_1) \right) (b_2 r_5 i_3)(b_2 r_2 i_2) b_1 r_3 i_2 \right] r_9 (n_1 r_8 n_3) \right\} r_9 s_2 \quad (1)$$

Такое формальное описание угрозы позволяет выделить признаки, по которым можно осуществлять поиск в автоматическом режиме записи о данной угрозе в базе данных, при сборе в сети Internet информации в интересах статистического анализа реализации угроз и др.

3. Модификация языка семантических сетей в интересах его использования для логико-лингвистического моделирования угроз

Семантические сети (СС) представляют собой следующий шаг в развитии машинно-читаемых представлений информации. Они дают возможность получать логические выводы на основании введенных правил.

6 Модальность (от лат. Modus – мера, способ) – оценка связи, устанавливаемой в высказывании, данная с той или иной точки зрения. Модальная оценка выражается с помощью модальных понятий: «необходимо», «возможно», «случайно», «доказуемо», «опровержимо», «обязательно», «разрешено», «хорошо» и т.п.

СС, как и СЦ, формально определяется тремя классами термов [5, 6]: 1) понятиями $X = \{X_1, X_2, ..., X_s\}$; 2) именами $I = \{i_1, i_2, ..., i_m\}$; 3) предикатами (типами отношений) $R = \{\rho, R_1, R_2, ..., R_q\}$, при этом предикат является особым, означающим «иметь имя», между понятиями в виде триплета ($x_k - r_j - x_s$) и между именами в виде триплета ($i_p - r_g - i_h$).

Запись на языке СС значительно богаче, так как с ее помощью можно выразить сложные совокупности фактов, описывать не только постоянные отношения, присущие объекту моделирования, но и ситуативные отношения, а также обобщать ситуации за счет выделения типовых структур отношений [6]. Это позволяет использовать их при создании иерархических баз знаний и разработке формальных процедур ситуационного поиска решений. Применительно к задачам ТЗИ этот язык пока непосредственно не использовался. В отличие от традиционного языка СС, применяемого чаще всего для поиска семантической информации [5, 6, 8], созданный затем язык ситуационного управления (ЯСУ) [16] применяется для описания порядка действий в ходе управления процессами и объектами, распознавания ситуаций и др. Вместе с тем, при описании угроз с использованием реляционных языков, в том числе языков СС и ЯСУ, не учитывается фактор времени, то есть не описывается динамика процессов реализации угроз, для чего, в основном, применяются аппараты марковских, полумарковских процессов и составных сетей Петри-Маркова [16]. Однако для построения математических моделей на основе указанных аппаратов, оценивания возможностей реализации угроз, определения состава актуальных угроз и т.д. требуются соответствующие исходные данные об угрозах, при этом из-за большого количества угроз и объема сведений о них для формирования исходных данных необходимо не только автоматизировать их поиск в базе знаний, но и подготавливать эти данные в нужном для расчетов формате. Для этого также может быть использовано описание угроз на основе реляционных языков.

Специфика задач ТЗИ обусловливает то, что, как и в интересах ситуационного управления был преобразован язык СС в ЯСУ [3], для решения задач ТЗИ язык СС тоже должен быть соответствующим образом изменен в части понятий, имен и отношений. Такими изменениями в языке СС могут быть, например, по аналогии с [6], отношения, устанавливаемые между действиями, соответствующими процессам реализации угроз, то есть отношения между предикатами, определяющими такие действия. В этом случае все предикаты, определяющие отношения, разбиваются на три подмножества: отношения между понятиями, между именами и отношения между предикатами, сами определяющие другие отношения.

Примеры некоторых понятий, имен и отношений между именами применительно к задачам описания и анализа угроз безопасности информации в ИС приведены в таблицах (табл.2 и 3).

Это позволяет в формальном описании угрозы связывать между собой последовательно выполняемые и обусловленные действия. Для описания угроз в этом случае используются четыре класса термов: 1) понятия $\{X_s\}$; 2) имена $\{i_m\}$; 3) предикаты, определяющие отношения между понятиями в виде триплетов ($x_k - r_j - x_s$) и между именами – ($i_p - r_g - i_h$); 4) предикаты, определяющие отношения между выполняемыми действиями в виде триплетов ($r_u - r_d - r_b$), к которым относятся, например, отношения следования, временные, каузативные и др. Семантическую сеть представляют часто в виде ориентированного графа, вершинами которого являются понятия и имена, а дугами – отношения между ними. Пример графа, описывающего понятия угрозы копирования файла конфиденциальной информации с использованием файла cookie без раскрытия сценария реализации, приведен на рис.2.

Рассмотрим отдельно описание способа (сценария) реализации угрозы копирования нужной нарушителю информации с использованием файла cookie.

Пример 3. Введем с учетом табл. 2 и 3 следующие обозначения: i_{21} – имя способа реализации угрозы копирования информации с использованием файла cookie; i_{12} – файл cookie; i_{16} – компьютер атакующего; i_{17} – атакуемый компьютер; i_{19} – компьютер со снифером для перехвата файлов cookies; i_{20} – поисковая система; $r_{21}^{(res)}$ – передача запроса на информационный ресурс в поисковую систему; $r_{21}^{(cook)}$ – передача файла cookie с поисковой машины на атакуемый хост; r_{25} – перехват файла cookie на компьютер-посредник с установленной программой «сниффер» и передача его на компьютер атакующего; r_{27} – внедрение в файл cookie вредоносного скрипта; $r_{21}^{(scr)}$ – передача файла cookie, инфицированного вредоносным скриптом, на атакуемый компьютер; $r_{20}^{(scr)}$ – прием и просмотр веб-браузером атакуемого компьютера файла cookie и запуск на исполнение внедренного в него скрипта; r_{24} – копирование файла; $r_{21}^{(f)}$ – передача скопированного вредоносным скриптом файла на хост атакующего.

Тогда формальная запись способа реализации угрозы на языке СС имеет вид:

Таблица 2

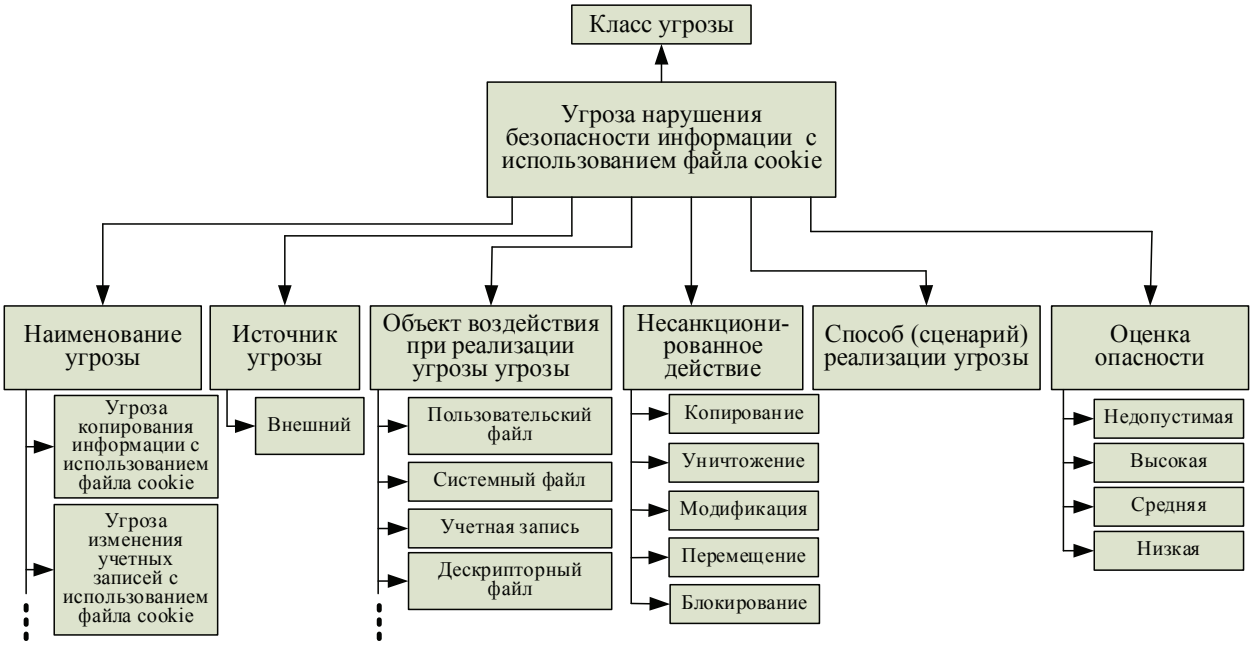
Примеры некоторых понятий и имен, используемых для описания угроз безопасности информации с применением семантических сетей

№	Понятие	Обозначение понятия	Имя понятия	Обозначение имени понятия
1	Наименование угрозы	X_0	Угроза копирования файла пользовательских данных с использованием файла cookie	i_1
2			Угроза уничтожения файла	i_2
3			Угроза подмены (модификации) файла	i_3
4			Угроза блокирования исполняемого файла	i_4
5	Класс угрозы	X_1	Угроза проникновения в операционную среду	i_5
6			Угроза отказа в обслуживании	i_6
7			Угроза перехвата трафика	i_7
8	Источник угрозы	X_2	Внешний	i_8
9			Внутренний	i_9
10	Объект воздействия	X_3	Каталог	i_{10}
11			Файл системный	i_{11}
12			Файл cookie	i_{12}
13			Текстовый файл	i_{13}
14			Графический файл	i_{14}
15			Исполняемый файл прикладной программы	i_{15}
16	Хост	X_4	Хост атакующего	i_{16}
17			Хост атакуемого	i_{17}
18			Хост-«жертва»	i_{18}
19			Хост – «посредник»	i_{19}
20			Поисковая система	i_{20}
21	Способ реализации	X_5	Идентификатор №1	i_{21}
22			Идентификатор №2	i_{22}
23	Состояние	X_6	Нормальное	i_{23}
24			Отказ в обслуживании	i_{24}
25			Угроза реализована	i_{25}
26			Угроза блокирована	i_{26}
27	Вспомогательное	X_7	Сетевой адрес	i_{27}
28			Предельное количество записей в буфер	i_{28}
29			Вредоносная программа	i_{29}
30			Количество отправленных пакетов	i_{30}
31	Уровень опасности угрозы	X_8	Недопустимая	i_{31}
32			Высокая опасность	i_{32}
33			Средняя опасность	i_{33}
34			Низкая опасность	i_{34}

Таблица 3

Примеры некоторых отношений между понятиями и именами при формальной записи процессов реализации угроз

№	Тип отношения	Содержание отношения	Обозначение
1	Временные	Быть одновременно	r_1
2		Быть позже	r_2
3		Быть раньше	r_3
4	Структурные	Быть в одном сегменте	r_4
5		Находиться за пределами сети	r_5
6		Содержаться в файле	r_6
7		Быть за межсетевым экраном	r_7
8	Сравнительные	Быть равным	r_8
9		Быть больше	r_9
10		Быть меньше	r_{10}
11	Классификационные	Содержать в качестве элемента	r_{11}
12		Относиться к чему либо (например, к категории – виду, классу, типу)	r_{12}
13	Идентифицирующие	Иметь имя	r_{13}
14		Быть чем либо (например источником, уязвимостью и т.д.)	r_{14}
15		Быть актуальной	r_{15}
16	Каузативные	Быть целью для...	r_{16}
17		Быть причиной для...	r_{17}
18		Быть следствием для...	r_{18}
19		Быть условием для...	r_{19}
20	Динамические	Принять что-то (пакет, сообщение, квитанцию и т.д.	r_{20}
21		Передать на	r_{21}
22		Повторить	r_{22}
23		Блокировать	r_{23}
24		Копировать	r_{24}
25		Перехватить	r_{25}
26		Ограничить	r_{26}
27		Модифицировать	r_{27}
28	Прагматические	Служить для	r_{28}
29		Находиться в состоянии	r_{29}
30		Обеспечить	r_{30}
31		Быть дополнением	r_{31}
32	Следования	Происходить непосредственно перед началом следующего действия	r_{32}
33		Происходить непосредственно после завершения предыдущего действия	r_{33}
34		Происходить параллельно с другим действием (другими действиями)	r_{34}



Формальная запись угрозы с использованием языка семантической сети без раскрытия способа ее реализации

$$(X_0\rho i_1)\&\left[(X_0r_{12}X_1)(X_1r_{13}i_5)\right]\&\left[(X_0r_{11}X_2)(X_2r_{13}i_8)\right]\&\left[(X_0r_{11}X_3)(X_3r_{13}i_{13})\right]\&$$
$$\&\left[(X_0r_{11}X_4)(X_4r_{13}i_{16})\right]\&\left[(X_0r_{11}X_4)(X_4r_{13}i_{17})\right]\&\left[(X_0r_{11}X_4)(X_4r_{13}i_{19})\right]\&$$
$$\&\left[(X_0r_{11}X_5)(X_5r_{13}i_{21})\right]\&\left[(X_0r_{11}X_6)(X_6r_{13}i_{25})\right]\&\left[(X_0r_{11}X_8)(X_8r_{13}i_{32})\right]$$

Рис.2. Семантическая сеть, описывающая понятие угрозы копирования файла пользовательских данных, без раскрытия способа ее реализации

$$X_5\rho i_{21}\ \&\ i_{17}r_{21}^{(res)}i_{20}\ \&\ i_{20}r_{21}^{(cook)}i_{17}\ \&\ i_{19}r_{25}i_{16}\ \&\ i_{16}r_{27}i$$
$$r_{27}i_{12}\ \&\ i_{16}r_{21}^{(scr)}i_{17}\ \&\ i_{17}r_{20}^{(scr)}i_{12}\ \& \qquad (2)$$
$$\&i_{17}r_{11}i_{13}\ \&\ i_{29}r_{24}i_{13}\ \&\ i_{17}r_{21}^{(f)}i_{16}$$

Подобные записи можно использовать в перспективных экспертных системах, предназначенных для автоматического или автоматизированного анализа угроз безопасности информации, в специализированных системах поддержки принятия решений по ТЗИ в ИС, при решении задач сбора и обработки данных об угрозах безопасности информации по результатам мониторинга сети Internet в интересах ведения Банка данных угроз безопасности информации ФСТЭК России и др. Расширение рассмотренных языков для логико-лингвистического моделирования угроз позволит, кроме того, формально описывать способы применения мер и средств защиты от актуальных угроз безопасности информации в ИС, способы обхода (преодоления) мер и средств защиты, а также подготавливать в автоматизированном режиме данные по угрозам для расчетов показателей оценки эффективности применения мер и средств защиты с использованием соответствующих математических моделей.

4. Выводы

1. Переход от вербальных к формальным описаниям угроз безопасности информации и необходимость использования элементов искусственного интеллекта для их анализа востребованы практикой организации и ведения ТЗИ от НСД, что обусловлено большим объемом данных об угрозах, обработка которых крайне затруднительна без ее автоматизации. Для формализации описаний угроз, то есть разработки их логико-лингвистических моделей, могут быть использованы реляционные языки описания. Логико-лингвистические модели угроз, подлежащие включению в соответствующие базы знаний, востребованы сегодня для обеспечения решения таких задач ТЗИ от НСД, как ведения Банка данных угроз безопасности информации ФСТЭК России, выявление состава актуальных угроз безопасности информации в различных классах ИС, формирование их описаний для включения в состав частных моделей угроз, автоматизированная подготовка данных для количественных оценок возможно-

стей их реализации, построение специализированных экспертных систем (систем поддержки принятия решений), необходимых для подготовки и принятия решений по ТЗИ от НСД, и др.

2. Для логико-лингвистического моделирования угроз безопасности информации могут использоваться языки RX-кодов, синтагматических цепей и семантических сетей. Наибольшими выразительными возможностями обладает язык семантических сетей, однако для несложных описаний угроз могут быть успешно применены также языки RX-кодов и синтагматических цепей. В связи с этим целесообразно создание на основе понятий, имен и отношений указанных языков единого реляционного языка, позволяющего не только на формальной основе описывать угрозы безопасности информации и процессы их реализации, но и применение мер и средств защиты от них, способы их обхода и преодоления, а также хранить результаты анализа и количественные оценки возможностей реализации каждой из актуальных угроз и эффективности мер и средств защиты от них в различных ИС.

Литература

1. Язов Ю.К., Соловьев С.В. Организация защиты информации в информационных системах от несанкционированного доступа. Монография. – Воронеж: Кварта. 2018. – 588 с. ISBN 978-5-93737-158-4
2. Киреева Н.В., Поздняк И.С., Филиппов Н.В. Подход к созданию экспертной системы оценки информационной безопасности телекоммуникационных систем // Электросвязь. 2022. №2. С. 61 – 66. ISSN 0013-5771
3. Зегжда П.Д., Анисимов В.Г., Супрун А.Ф., Анисимов Е.Г., Сауренко Т.Н. Модели и метод поддержки принятия решений по обеспечению информационной безопасности информационно-управляющих систем // Проблемы информационной безопасности. Компьютерные системы. 2018 . №1. С. 43 – 47. ISSN 2071-8217
4. Букатова И.Л. Эволюционное моделирование: идеи, основы теории, приложения. М.: Знание, 2020. - 888 с.
5. Балан В.П., Душкин А.В., Новосельцев В.И., Сумин В.И. Введение в системное проектирование интеллектуальных баз знаний / Под ред. В.И. Новосельцева – М.: Горячая линия – Телеком. 2016. – 107 с. ISBN 978-5-9912-0589-4
6. Кравченко Ю.А. Задачи семантического поиска, классификации, структуризации и интеграции информации в контексте проблем управления знаниями // Известия ЮФУ. Технические науки. 2016. №7(180). С. 6 – 18. ISSN 1999 – 9429
7. Марченко А.А. Метод автоматического построения онтологических баз знаний. II. Автоматическое определение семантических отношений в онтологической сети // Кибернетика и системный анализ. 2016. Том 52, №2. ISSN 0023-1274
8. Ярушкина Н.Г., Мошкин В.С., Филиппов А.А., Гуськов Г.Ю. Романов А.А., Наместиков А.М. Разработка программной системы семантического анализа контента социальных медиа. Математическое моделирование инфокоммуникационных систем // Радиотехника. 2018. №6. С.73 – 79. ISSN 033-8486
9. Umara Noor, Zahid Anwar, Asad Waqar Malik, Sharifullah Khan, Shahzad Saleem A machine learning framework for investigating data breaches based on semantic analysis of adversary's attack patterns in threat intelligence repositories // Future Generation Computer Systems. 2019. Volume 95. C.467 – 487. DOI 10.1016/Future.2019.01.022
10. Дойникова Е.В. Федорченко А.В., Котенко И.В., Новикова Е.С. Методика оценивания защищенности на основе семантической модели метрик и данных // Вопросы кибербезопасности. 2021. №1 (41). С. 29 – 40. DOI 10.21681/ 2311-3456-2021-1-29-40
11. Васильев В.И., Вульфин А.М., Кучкарова Н.В. Автоматизация анализа уязвимостей программного обеспечения на основе технологии text mining // Вопросы кибербезопасности. 2020. №4 (38). С. 22 – 31. DOI 10.21681/ 2311-3456-2020-04-22-31
12. Федорченко А.В., Дойникова Е.В., Котенко И.В. Автоматизированное определение активов и оценка их критичности для анализа защищенности информационных систем // Труды СПИИРАН. 2019. Т.18. №5. С. 1182-1211. DOI 10.15622/ sp.2019.18.5.1182-1211
13. Гаршина В.В., Степанцов В.А., Данковцева А.Ю. Семантический анализ информационных рисков и угроз на основе онтологии стандарта ISO/IES 27001 // Вестник Воронежского государственного университета, серия «Системный анализ и информационные технологии». 2018. №4. С. 73 – 80. ISSN 1995 – 5499
14. Бубакар И., Будько М.Б., Будько М.Ю., Гирик А.В. Онтологическое обеспечение управления рисками информационной безопасности. Труды ИСП РАН. 2021,том 33, вып. 5. С. 41-64. DOI: 10.15514/ISPRAS–2021–33(5)–3
15. Васильев В. И., Вульфин А. М., Кириллова А. Д., Кучкарова Н. В. Методика оценки актуальных угроз и уязвимостей на основе технологий когнитивного моделирования и Text Mining // Системы управления, связи и безопасности. 2021. № 3. С. 110-134. DOI: 10.24412/2410-9916-2021-3-110-134
16. Язов Ю.К., Анищенко А.В. Сети Петри-Маркова и их применение для моделирования процессов реализации угроз безопасности информации в информационных системах. Монография. – Воронеж: Кварта, 2020. – 173 с. ISBN 978-5-93737-187-4

LOGICAL-LINGUISTIC MODELING OF SECURITY THREATS INFORMATION IN INFORMATION SYSTEMS

Yazov Yu.K.[7], Soloviev S.V.[8], Tarelkin M.A.[9]

Purpose: *assessment of the possibility, definition of conditions and a brief description of the relational languages of logical-linguistic modeling for a formalized description and presentation of the processes of implementing information security threats in information systems.*

Method: *application of the logical-linguistic modeling apparatus, which makes it possible to formally describe information security threats and a set of actions performed in the course of their implementation, taking into account the capabilities of relational description languages, such as Codd's language, context-free plex-language, RX-code language, syntagmatic chains and semantic networks.*

Result: *a brief description and comparative analysis of relational description languages and features that affect the possibility of their use for describing threats to information security and logical-linguistic modeling of their implementation processes are given. The expediency of such modeling is shown when creating promising expert systems designed for automated and automatic analysis of threats, when maintaining a data bank of threats based on the results of monitoring publications about them on the Internet.*

Examples of constructing formal logical-linguistic descriptions of well-known threats of computer attacks on information systems using RX-code languages and semantic networks are given, proposals are made for expanding the language of semantic networks to describe threats, taking into account new data on threats and methods for their implementation.

It is noted that the proposed approach to modeling the processes of implementation of information security threats, as a rule, is applicable in the absence of the need to take into account the time factor when assessing the possibilities of their implementation.

Keywords: *security threat, relational language, assessment, functional model, Petri-Markov net, security measure.*

References

1. Jazov Ju.K., Solov'ev S.V. Organizacija zashhity informacii v informacionnyh sistemah ot nesankcionirovannogo dostupa. Monografija. – Voronezh: Kvarta. 2018. – 588 s. ISBN 978-5-93737-158-4
2. Kireeva N.V., Pozdnjak I.S., Filippov N.V. Podhod k sozdaniju jekspertnoj sistemy ocenki informacionnoj bezopasnosti telekommunikacionnyh sistem // Jelektrosvjaz'. 2022. №2. S. 61 – 66. ISSN 0013-5771
3. Zegzhda P.D., Anisimov V.G., Suprun A.F., Anisimov E.G., Saurenko T.N. Modeli i metod podderzhki prinjatija reshenij po obespecheniju informacionnoj bezopasnosti informacionno-upravljajushhih sistem // Problemy informacionnoj bezopasnosti. Komp'juternye sistemy. 2018 . №1. S. 43 – 47. ISSN 2071-8217
4. Bukatova I.L. Jevoljucionnoe modelirovanie: idei, osnovy teorii, prilozhenija. M.: Znanie, 2020. – 888 c.
5. Balan V.P., Dushkin A.V., Novosel'cev V.I., Sumin V.I. Vvedenie v sistemnoe proektirovanie intellektual'nyh baz znanij / Pod red. V.I. Novosel'ceva – M.: Gorjachaja linija – Telekom. 2016. – 107 s. ISBN 978-5-9912-0589-4
6. Kravchenko Ju.A. Zadachi semanticheskogo poiska, klassifikacii, strukturizacii i integracii informacii v kontekste problem upravlenija znanijami // Izvestija JuFU. Tehnicheskie nauki. 2016. №7(180). S. 6 – 18. ISSN 1999 – 9429
7. Marchenko A.A. Metod avtomaticheskogo postroenija ontologicheskih baz znanij. II. Avtomaticheskoe opredelenie semanticheskih otnoshenij v ontologicheskoj seti // Kibernetika i sistemnyj analiz. 2016. Tom 52, №2. ISSN 0023-1274
8. Jarushkina N.G., Moshkin V.S., Filippov A.A., Gus'kov G.Ju. Romanov A.A., Namestikov A.M. Razrabotka programmnoj sistemy semanticheskogo analiza kontenta social'nyh media. Matematicheskoe modelirovanie infokommunikacionnyh sistem // Radiotehnika. 2018. №6. S.73 – 79. ISSN 033-8486
9. Umara Noor, Zahid Anwar, Asad Waqar Malik, Sharifullah Khan, Shahzad Saleem A machine learning framework for investigating data breaches based on semantic analysis of adversary's attack patterns in threat intelligence repositories // Future Generation Computer Systems. 2019. Volume 95. S.467 – 487. DOI 10.1016/Future.2019.01.022

7 Yury K. Yazov, Dr.Sc., Professor, Chief Researcher of the Department of FAU «GNIII PTZI FSTEC of Russia», Voronezh, Russia. E-mail: Yazoff_1946@mail.ru.

8 Sergey V. Soloviev, Head of Department of the FAU «GNII PTZI FSTEC of Russia», Voronezh, Russia. E-mail: gniii@fstec.ru.

9 Mikhail A. Tarelkin, Senior Researcher, FAU «GNII PTZI FSTEC of Russia», Voronezh, Russia. E-mail: gniii@fstec.ru.

10. Dojnikova E.V. Fedorchenko A.V., Kotenko I.V., Novikova E.S. Metodika ocenivanija zashhishhennosti na osnove semanticheskoj modeli metrik i dannyh // Voprosy kiberbezopasnosti. 2021. №1 (41). S. 29 – 40. DOI 10.21681/ 2311-3456-2021-1-29-40
11. Vasil'ev V.I., Vul'fin A.M., Kuchkarova N.V. Avtomatizacija analiza ujazvimostej programmnogo obespechenija na osnove tehnologii text mining // Voprosy kiberbezopasnosti. 2020. №4 (38). S. 22 – 31. DOI 10.21681/ 2311-3456-2020-04-22-31
12. Fedorchenko A.V., Dojnikova E.V., Kotenko I.V. Avtomatizirovannoe opredelenie aktivov i ocenka ih kritichnosti dlja analiza zashhishhennosti informacionnyh sistem // Trudy SPIIRAN. 2019. T.18. №5. S. 1182-1211. DOI 10.15622/ sp.2019.18.5.1182-1211
13. Garshina V.V., Stepancov V.A., Dankovceva A.Ju. Semanticheskij analiz informacionnyh riskov i ugroz na osnove ontologii standarta ISO/IES 27001 // Vestnik Voronezhskogo gosudarstvennogo universiteta, serija «Sistemnyj analiz i informacionnye tehnologii». 2018. №4. S. 73 – 80. ISSN 1995 – 5499
14. Bubakar I., Bud'ko M.B., Bud'ko M.Ju., Girik A.V. Ontologicheskoe obespechenie upravlenija riskami informacionnoj bezopasnosti. Trudy ISP RAN. 2021,tom 33, vyp. 5. S. 41-64. DOI: 10.15514/ISPRAS–2021–33(5)–3
15. Vasil'ev V. I., Vul'fin A. M., Kirillova A. D., Kuchkarova N. V. Metodika ocenki aktual'nyh ugroz i ujazvimostej na osnove tehnologij kognitivnogo modelirovanija i Text Mining // Sistemy upravlenija, svjazi i bezopasnosti. 2021. № 3. S. 110-134. DOI: 10.24412/2410-9916-2021-3-110-134
16. Jazov Ju.K., Anishhenko A.V. Seti Petri-Markova i ih primenenie dlja modelirovanija processov realizacii ugroz bezopasnosti informacii v informacionnyh sistemah. Monografija. – Voronezh: Kvarta, 2020. – 173 s. ISBN 978-5-93737-187-4