УДК 004.413.4

НЕКОТОРЫЕ ПОДХОДЫ К ОЦЕНКЕ ИНФОРМАЦИОННЫХ РИСКОВ С ИСПОЛЬЗОВАНИЕМ НЕЧЁТКИХ МНОЖЕСТВ

Минзов Анатолий Степанович [1], Шевяхов Максим Юрьевич [2]

[1] Доктор технических наук, профессор;
ГОУ ВПО Международный университет природы, общества и человека «Дубна»,
Институт системного анализа и управления;
141980, Московская обл., г. Дубна, ул. Университетская, 19;
e-mail: a@minzov.ru.

2 Магистр;
ГОУ ВПО Международный Университет природы, общества и человека «Дубна»,
Институт системного анализа и управления;
141980, Московская обл., г. Дубна, ул. Университетская, 19;
e-mail: maxim.cat@gmail.com.

В статье проводится анализ различных подходов к оценке информационных рисков и предлагается новый подход с использованием теории нечётких множеств.

Ключевые слова: информационный риск, нечёткие множества, оценка риска.

SOME APPROACHES TO AN ESTIMATION OF INFORMATION RISKS WITH USE OF FUZZY SETS

Minzov Anatoliy[1], SHevyakhov Maxim[2]

[1]Doctor of Science in Engineering, Professor;
Dubna International University of Nature, Society, and Man,
Institute of system analysis and management;
141980, Dubna, Moscow reg., Universitetskaya str., 19;
e-mail: a@minzov.ru.

[2]Magistr;
Dubna International University of Nature, Society, and Man,
Institute of system analysis and management;
141980, Dubna, Moscow reg., Universitetskaya str., 19;
e-mail: maxim.cat@gmail.com.

The article carries out the analysis of various approaches to an estimation of information risks. The new approach with use of the theory of fuzzy sets is offered.

Keywords: information risk, fuzzy sets, a risk estimation.

Введение

Информационная безопасность в настоящее время становится необходимым условием успешного развития хозяйствующего субъекта. Риск компрометации информации влияет на материальные и нематериальные активы организации и, в конечном счёте, на результаты её производственно-экономической деятельности. В связи с большим числом информационных рисков, широким разбросом значений ущерба при их реализации и ограниченностью бюджета на информационную безопасность хозяйствующего субъекта возникает задача рационального финансирования затрат на защиту информации. Возможна и другая постановка задачи – при фиксированном объеме финансовых вложений необходимо снизить уровень риска компрометации информации на максимальную величину.

В настоящее время оценка информационных рисков проводится методами, требующими статистических данных по инцидентам, либо использующими некоторые категории опасности информационных рисков. Недостатком таких методов является тот факт, что информационные риски

имеют чаще всего субъективные значения, что вносит существенную погрешность в результаты их оценки. С другой стороны, оценка рисков с помощью экспертных методов вносит помеху в виде неточности экспертной оценки.

Последним фактом, затрудняющим принятие решения по обработке выявленного риска – малое количество параметров, которыми он характеризуется. Большинство оценок информационного риска опираются на два показателя – величину возможного ущерба риска и вероятность его возникновения. Однако очень часто требуется также знание величины затрат на снижение риска до приемлемого уровня. Иногда могут использоваться и другие показатели: характер риска (периодический, случайный или однократный), метод финансирования программ по снижению риска (единовременные или регулярные затраты на уменьшение риска). Введение всё новых и новых элементов усложняет модель оценки риска и создает определенные трудности её практического использования.

В работе рассматриваются модели оценки информационных рисков на основе теории нечётких множеств.

Информационный риск и его оценка

Стандарт информационной безопасности [1] определяет оценку информационного риска как комплексную оценку двух показателей:

- возможного ущерба, наносимого компании при нарушении информационной безопасности;
- вероятности наступления такого нарушения.

Оба показателя не всегда могут быть точно численно определены, поэтому возможным решением является проведение их оценок по нечётким данным. Такими данными могут быть мнения экспертов. Далее оценки передаются специалисту по информационной безопасности, который решает, какие риски следует уменьшить, а какие нет. В доминирующей методике оценки информационных рисков уровню риска присваивается произведение значений величины возможного ущерба и вероятности возникновения риска. Далее из множества рисков выбирается тот, уровень которого наибольший, и он устраняется. Следом за ним устраняется второй по уровню риск, далее третий и т.д. Оценку риска по такой модели можно интерпретировать как средняя ожидаемая потеря от риска за расчётный промежуток времени. Эта интерпретация весьма груба. В самом деле, если риск случится, то мы получим ущерб от риска полностью, а не частично.

Существенным недостатком такого подхода является отсутствие величины затрат на снижение риска при принятии решения о судьбе риска. В самом деле, самый высокий по уровню риск может потребовать столько затрат, сколько в сумме потребуют четыре следующих за ним риска. Видно, что уменьшить самый высокий по уровню риск менее выгодно, нежели уменьшить четыре последующих за ним риска. Следует также помнить тот факт, что есть зависимые друг от друга риски. Зависимость может быть двух видов: появление первого увеличивает вероятность появления второго, появление первого уменьшает вероятность появления второго. Также зависимость может проявляться и при минимизации рисков: устраняя один из рисков, мы снизим вероятность появления второго. Расчёт подобных зависимостей довольно сложен и должен проводиться экспертом по информационной безопасности.

Введение в качестве новой переменной стоимости снижения риска до приемлемого уровня, несомненно, усложнит систему принятия решений. К тому же, часто встречаемой ситуацией является риск, по которому нет полной статистики. Если расширять описанную выше двухфакторную модель новыми параметрами, то необходимо будет получить точную формулу для оценки уровня риска и принятия решения по риску. А в случае с неточными данными непонятно, какое значение параметров подставлять в формулу. В обоих случаях можно с точно определённых переменных и формул перейти на качественные оценки переменных и правила нечёткого вывода типа «если А, то В». Для этих целей хорошо подходит теория нечётких множеств, применение которой упрощает принятие решения по сложно формализованным и неформализованным задачам.

Теория нечётких множеств

Понятие «нечёткое множество» введено Л. А. Заде в 1965 г. [2]. В то же время он выпустил первые труды по теории нечётких множеств. Теория представляет собой аппарат формализации неопределённостей, возникающих при моделировании реальных объектов, а именно, при использовании естественных слов для описания объекта.

Рассмотрим U – универсальное множество, к которому относятся все результаты наблюдений. Чёткое, канторовское подмножество A в U можно определить как набор пар из значения $u \in U$ и степени принадлежности к множеству $\mu_A(u)$ этого значения. При этом степень принадлежности принимает только два значения: 0, «не принадлежит», и 1, «принадлежит». Если разрешить $\mu_A(u)$ принимать и промежуточные значения, то получится множество A, про элементы которого нельзя точно сказать, принадлежат они множеству A или нет. Такое множество называется нечётким. В этом смысле чёткое множество является частным случаем нечёткого множества.

По аналогии с канторовскими множествами были введены логические и алгебраические операции для нечётких множеств, а также выведены свойства, подобные свойствам чётких множеств [3]. Например, коммутативность, ассоциативность и теоремы де Моргана. С помощью всего вышеперечисленного мы получаем минимальный математический аппарат, необходимый для работы с нечёткими множествами.

Следующими двумя важными понятиями являются нечёткая и лингвистическая переменная.

Нечёткая переменная [3] характеризуется тройкой $\langle \alpha, X, A \rangle$, где

α – наименование переменной;

X – универсальное множество (область определения α);

A – нечёткое множество на X, описывающее ограничения на значения нечёткой переменной α (то есть $\mu_A(u)$).

Лингвистическая переменная [3] определяется следующим образом: $\Omega = \langle \omega, T(\omega), U, G, M \rangle$, где

ω – наименование лингвистической переменной;

T – множество её значений, представляющих собой наименования нечётких переменных, областью определения которых является множество U. Множество T называется базовым терм-множеством лингвистической переменной;

G – синтаксическая процедура, позволяющая оперировать элементами терм-множества T, в частности, генерировать новые термы;

M – семантическая процедура, позволяющая превратить каждое новое значение лингвистической переменной, образуемое процедурой G в нечёткую переменную, то есть сформировать соответствующее нечёткое множество.

Лингвистическая переменная служит для качественного описания какого-либо показателя, сопоставляя естественному, словесному описанию нечёткое множество. Такой подход позволяет высказывать суждения о значении показателя с помощью качественных оценок. К примеру, о вероятности возникновения угрозы можно сказать «высокая» вместо «78%». С другой стороны, вероятность 78% может быть и низкой, и средней. Знание качественной оценки зачастую важнее знания количественной оценки, что ставит использование теории нечётких множеств на первый план.

Среди нечётких переменных особое положение занимают нечёткие числа и нечёткие интервалы. Нечёткая величина – нечёткая переменная, определенная на множестве действительных чисел [3].

В общем случае нечётким интервалом называется выпуклая нечёткая величина [3]. Нечётким числом называется выпуклая унимодальная нечёткая величина [3].

Нечёткое число и нечёткий интервал позволяют описать такие термы лингвистической переменной, как «около 150», «где-то между 56 и 60», «гораздо больше 100». В общем случае нечёткое число и интервал не являются нормальными, но в большинстве случаев это подразумевается.

Мы нашли инструмент, позволяющий оперировать нечёткими состояниями. Попробуем описать информационный риск с помощью этого инструмента.

Особенности применения теории нечётких множеств

Для начала опишем существующую двухфакторную модель оценки риска в терминах нечётких множеств. Используем для этого стандарт NIST 800-30 [6]. Для описания вероятности возникновения возьмём лингвистическую переменную «P» с тремя термами: «high», «middle» и «low». Для описания величины ущерба от риска возьмём лингвистическую переменную «U» с тремя термами: «high», «middle» и «low». Для описания уровня риска возьмём лингвистическую переменную «RiskLevel» с тремя термами: «high», «middle», «low».

По стандарту NIST 800-30 уровень риска определяется как наименьший из уровней ущерба и вероятности риска [6]. Таким образом, получим следующий набор правил [6]:

1) Если «P» = «low» и «U» = «low», то «RiskLevel» = «low».
2) Если «P» = «low» и «U» = «middle», то «RiskLevel» = «low».
3) Если «P» = «low» и «U» = «high», то «RiskLevel» = «low».
4) Если «P» = «middle» и «U» = «low», то «RiskLevel» = «low».
5) Если «P» = «middle» и «U» = «middle», то «RiskLevel» = «middle».
6) Если «P» = «middle» и «U» = «high», то «RiskLevel» = «middle».
7) Если «P» = «high» и «U» = «low», то «RiskLevel» = «low».
8) Если «P» = «high» и «U» = «middle», то «RiskLevel» = «middle».
9) Если «P» = «high» и «U» = «high», то «RiskLevel» = «high».

IF		THEN	
P	U	DoS	RiskLevel
low	low	1.00	low
low	medium	1.00	low
low	high	1.00	low
medium	low	1.00	low
medium	medium	1.00	medium
medium	high	1.00	medium
high	low	1.00	low
high	medium	1.00	medium
high	high	1.00	high

Рис. 1. Нечёткие правила вывода, заданные таблично в программе FuzzyTECH

Чтобы самостоятельно применить вышеприведённые правила на практике, необходимо проделать следующие шаги. Сначала необходимо показатели риска фаззифицировать, то есть вычислить степень принадлежность каждого из них к каждому терму из соответствующего терм-множества. Сделать это можно двумя путями. Либо мы даём группе экспертов оценить, к какому уровню относятся показатели конкретного риска, что даёт нам готовую степень принадлежности. Либо мы заранее задаём в модели численные показатели термов (то есть для каждого терма – функцию принадлежности), а экспертную группу просим оценить точное значение показателей риска. Далее мы эти оценки переводим в степень принадлежности.

Вторым шагом мы входим в цикл, в котором будем рассматривать каждое из наших правил. В данном примере их девять, но в зависимости от ситуации их может быть больше или меньше.

Для каждого правила необходимо проделать следующие операции. Сначала оцениваем истинность каждого из равенств. Значение истинности будет равно функции принадлежности соответствующей переменной к терму, указанному справа от знака «равно». Дальше необходимо оценить истинность правила. Она будет равна минимальному значению истинности каждого из равенств, вхо-

дящих в правило. Результатом применения каждого правила будет нечёткое множество, которое повторяет терм, указанный в качестве результата, но степень принадлежности у него равна истинности правила в тех случаях, когда первая выше второй.

После цикла мы получим нечёткие множества в количестве нечётких правил вывода. Итоговым результатом будет нечёткое множество, равное объединению всех полученных нечётких множеств.

Чтобы не проделывать все эти процедуры самостоятельно, можно воспользоваться программой FuzzyTECH. Данная программа позволяет оперировать лингвистическими переменными и создавать для них продукционные правила вывода. В интерактивном режиме можно наблюдать за показателями каждой из переменных, а также за уровнем истинности каждого из правил.

Сначала создадим две входных и одну выходную переменную, соответствующие лингвистическим переменным «P», «U» и «RiskLevel», соответственно. Потом создадим блок правил, в который занесём правила, описанные выше. Запустив интерактивный режим, можно изменять значения переменных «P» и «U», наблюдать за степенью истинности правил и наблюдать за результатом переменной «RiskLevel».

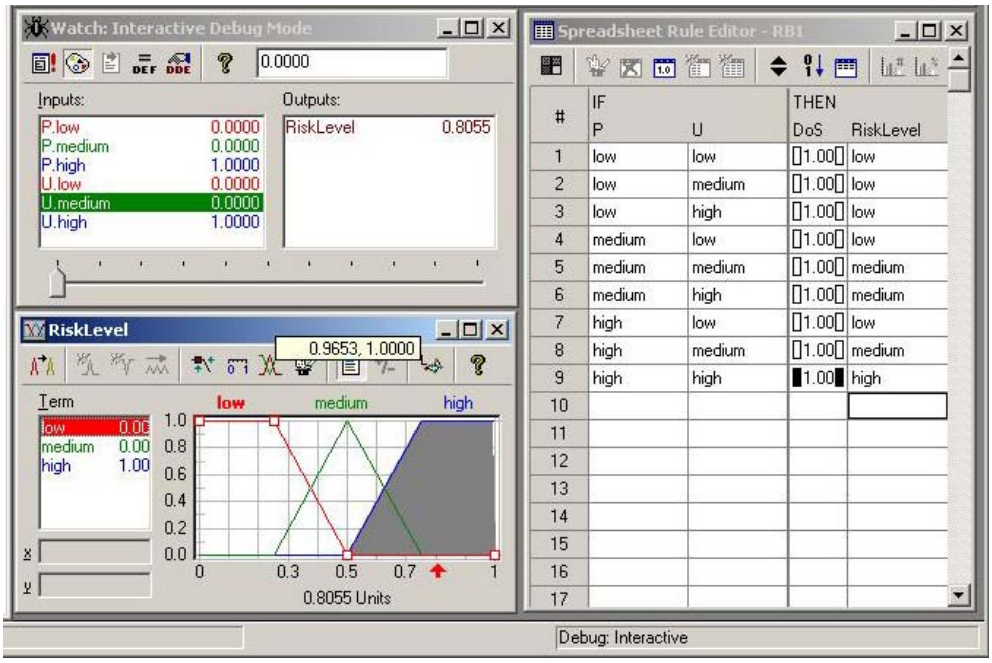

Рис. 2. Интерактивный режим программы FuzzyTech

Пользователь программы может предварительно выставить характерные значения для каждого из термов, либо пользоваться термами без численных значений для качественной оценки параметров.

На рис. 2 показана ситуация, когда пользователь для каждой переменной вводит вручную степень принадлежности к соответствующим термам. Полученные значения обрабатываются в соответствии с правилами, и на таблице в правой части рис. 2 отображается истинность правила в виде чёрного прямоугольника. Прямоугольник, закрашенный полностью, показывает на истинность, равную 1, прямоугольник не закрашенный – на истинность, равную 0. Промежуточным значениям соответствует прямоугольник, закрашенный частично. Отображение истинности правила позволяет следить за ошибками при переносе правил в программу, а также за влиянием каждого из правил на конечный результат.

Рассмотрим пример, показанный на рис. 2. Эксперт оценил показатели риска и получил, что вероятность возникновения риска соответствует терму «высокий», ущерб от риска также соответствует терму «высокий». Заметим, что эти термы похожи только по названию. Они относятся к разным терм-множествам («P» и «U), поэтому нечёткие величины, соответствующие этим термам, различны. Перебирая по очереди все правила, видим, что истинность всех правил, кроме одного, равняется нулю. Единственное правило с ненулевой истинностью – это последнее правило в таблице, которое гласит, что если «P» = «high» и «U» = «high», то «RiskLevel» = «high». В программе в таблице справа ему соответствует полностью закрашенный прямоугольник, а в окошке слева внизу показано значе-

ние переменной «RiskLevel» и каждого из её термов. Видим, что терму «high» соответствует степень принадлежности, равная 1.

Таким образом, в интерактивном режиме программы FuzzyTECH можно не только видеть конечный результат, но и следить за промежуточными операциями. Данная возможность необходима при внесении новых переменных и правил в процедуру оценки риска, демонстрация промежуточных результатов контролирует перенос правил нечёткого вывода в программу.

Добавление стоимости снижения риска

Модель, основанная на двух факторах риска, а именно вероятности возникновения и величины ущерба, является неполной. В ряде случаев необходимо знать затраты на снижение уровня риска путём уменьшения вероятности либо ущерба от риска. Если не принимать в расчёт эту величину, то можно потратить на устранение риска ресурсов больше, чем ожидаемый ущерб от риска.

В стандарте ГОСТ ИСО 17799-2005 [1] рекомендуется при оценке риска учитывать и отношение стоимости снижения риска к ущербу от риска. Если эта величина больше единицы, то такой риск рекомендуется принять. Если же это отношение меньше единицы, то имеет смысл устранить этот риск. Заметим, что окончательное решение принимается исходя не только из этого отношения, но и из величины уровня риска.

Для добавления этой переменной в оценку риска с помощью нечётких множеств необходимо ввести лингвистическую переменную, характеризующую стоимость снижения риска. Назовём её «Z», её термы «high», «middle», «low». Введём также выходную переменную «doing», показывающую то, насколько целесообразно уменьшить такой риск. У неё будет пять термов, которые мы назовём «very_positive», «positive», «zero», «negative», «very_negative». Следующим шагом введём правила нечёткого вывода для преобразования входных переменных «U» и «Z» в выходную переменную «doing». Заметим, что значения термов для «U» и «Z» должны определяться одинаковым образом, то есть, если какой-то величине ущерба соответствует терм «high», то такому же значению величины стоимости снижения риска соответствует терм «high». Это необходимое условие для сравнения величин ущерба и стоимости снижения риска.

Зададим нечёткие правила вывода в следующем виде:

1) Если «Z» = «high» и «U» = «high», то «doing» = «zero».
2) Если «Z» = «high» и «U» = «middle», то «doing» = «negative».
3) Если «Z» = «high» и «U» = «low», то «doing» = «very_negative».
4) Если «Z» = «middle» и «U» = «high», то «doing» = «positive».
5) Если «Z» = «middle» и «U» = «middle», то «doing» = «zero».
6) Если «Z» = «middle» и «U» = «low», то «doing» = «negative».
7) Если «Z» = «low» и «U» = «high», то «doing» = «very_positive».
8) Если «Z» = «low» и «U» = «middle», то «doing» = «positive».
9) Если «Z» = «low» и «U» = «low», то «doing» = «zero».

IF		THEN	
U	Z	DoS	doing
low	low	1.00	zero
low	medium	1.00	negative
low	high	1.00	very_negative
medium	low	1.00	positive
medium	medium	1.00	zero
medium	high	1.00	negative
high	low	1.00	very_positive
high	medium	1.00	positive
high	high	1.00	zero

Рис. 3. Нечёткие правила вывода, заданные таблично в программе FuzzyTECH

С помощью программы FuzzyTECH мы можем по оцененным значениям переменных «U» и «Z» получить значение переменной «doing». Если переменная принимает значения «positive» и «very_positive», то такой риск целесообразно устранить, потому что на снижение уровня риска мы потратим меньше ресурсов, чем потеряем при реализации риска. Если же значение переменной «doing» принимает значения «negative» и «very_negative», то нецелесообразно снижать уровень такого риска. Наличие приставки «very_» указывает на крайнее значение, когда переменные стоимости снижения риска и возможного ущерба от него принимают противоположные значения. В случае, когда переменная принимает значение «zero» о целесообразности снижения риска сложно судить, так как стоимость снижения риска и возможный ущерб от него принимают одно лингвистическое значение.

Таким образом, мы получаем две переменные, которые характеризуют уровень риска и целесообразность снижения уровня риска. Эти две переменные помогают как специалисту службы безопасности при составлении списка наиболее высоких рисков, так и лицу, принимающему решения при оценке целесообразности снижения рисков с точки зрения затрат. Таким образом, в зависимости от ситуации, могут быть разные алгоритмы выбора среди группы рисков тех, которые необходимо уменьшить. Если ограничено финансирование службы безопасности, то необходимо из группы рисков, самых целесообразных для снижения уровня, выбрать те, уровень риска в которых самый большой. Если есть ещё ресурсы, то выбрать риски в высокой степенью целесообразности и средним уровнем риска. В случае, когда требуется снизить общий уровень риска, первым устраняют риски из категории самых выгодных для снижения и с наибольшим уровнем риска. После этого устраняют менее выгодные риски и т.п. Примечательно, что две полученные переменные являются простыми для понимания и использования.

Вывод

Как показано выше, использование теории нечётких множеств может помочь при оценке риска в условиях неопределённости значений показателей риска, а также при выборе из группы рисков нескольких, которые должны удовлетворять определённому критерию. Использование программы FuzzyTECH помогает при оценке рисков. Недостатком применения нечётких множеств является субъективность оценки риска в нечётких терминах и субъективность правил вывода. В данной работе правила были составлены с учётом требований и рекомендаций стандартов Российской Федерации и США.

Список литературы

1. Стандарт ГОСТ Р ИСО/МЭК 17799-2005 «Информационная технология. Практические правила управления информационной безопасностью».
2. Zadeh L. A. Fuzzy sets. Information and Control. – 1965. – Vol. 8, № 3. – Pp. 338-353.
3. Круглов В. В., Дли М. И., Голунов Р. Ю. Нечёткая логика и искусственные нейронные сети. Учеб. пособие. – М.: Издательство Физико-математической литературы, 2001. – С. 224.
4. Леоненков А. В. Нечёткое моделирование в среде MATLAB и FuzzyTech. – СПб.: БХВ-Петербург, 2005. – С. 739.
5. Дюбуа Д., Прад А. Теория возможностей. Приложение к представлению знаний в информатике: пер. с фр. – М.: Радио и связь, 1990. – С. 288.
6. Risk Management Guide for Information Systems // NIST, Special publication 800-30.