

УПРАВЛЕНИЕ И ЛОГИКО-ВЕРОЯТНОСТНЫЕ МОДЕЛИ РИСКА С ГРУППАМИ НЕСОВМЕСТИМЫХ СОБЫТИЙ

Н.А. Махутов⁽¹⁾, Е.Д. Соложенцев⁽²⁾

⁽¹⁾ *Институт машиноведения РАН, г. Москва;*

⁽²⁾ *Институт проблем машиноведения РАН, г. Санкт-Петербург*

Рассмотрена связь статистических данных, управления, эффективности и риска на основе логико-вероятностной (ЛВ) теории риска с группами несовместных событий (ГНС). Выполнены систематизации и обобщения по ЛВ-моделям риска с ГНС в разных приложениях. Сделан акцент на логические основы теории риска, переход от базы данных к базе знаний. Описаны области приложений, достоинства, отличия, типы моделей и особенности ЛВ-теории риска с ГНС.

ВВЕДЕНИЕ

Вероятностная (В) логика Дж. фон Неймана и логико-вероятностное исчисление И.А. Рябинина для решения прямых задач риска появились в 1950—1960 гг. независимо друг от друга для разных приложений [1, 2]. В принципе, можно было использовать один термин, считая ЛВ-исчисление развитием В-логики. Однако ЛВ-исчисление породило целое научное направление и стало основой ЛВ-теории риска. Решение обратных задач риска с оценкой вероятностей по набору логических (Л) суждений впервые рассмотрел Нильс Нилссон, но аналитическое решение удалось получить только для простейших случаев. Обратные задачи риска без каких-либо ограничений успешно решаются в ЛВ-теории риска с группами несовместных событий (ГНС) [2, 3].

Основы ЛВ-теории риска с ГНС для целей управления создавались в течение 10 лет [2—4]. Она применялась для решения важных экономических проблем классификации (кредитные риски, рейтинги, мониторинг), инвестирования (портфель ценных бумаг), эффективности (управление социальными процессами), менеджмента (управление риском неуспеха по функциям, предметным об-

ластям, достижению целей и качеству функционирования), коррупции и взяток (выявление взяток по статистике параметров функционирования учреждения, поведения чиновников, обслуживания), управления развитием и испытаниями (организационные и экономические системы, сложные машины и технологии). В процессе решения выявились ее достоинства в сравнении с другими теориями: в два раза большая точность в распознавании «плохих» и «хороших» объектов; в семь раз большая робастность (стабильность) в распознавании объектов; абсолютная прозрачность в оценке и анализе риска; возможность решения новых задач анализа риска объектов; возможность управлять риском и эффективностью и др.

Логико-вероятностная теория риска с ГНС отличается видом используемой информации, типом связей между переменными, законами распределений переменных, методами решения обратных задач, использованием комбинированных моделей риска и др. (табл. 1).

ЛВ-теория риска с ГНС имеет следующие атрибуты: *классификатор научной области* — управление и эконометрика, *проблема* — управление риском в сложных системах, *идея* — введение в табличную БД конечных множеств, что позволяет



Таблица 1

Отличия ЛВ-теории риска с ГНС

В ЛВ-теории риска используются	В других теориях используются
Несколько уровней значений параметров	Два уровня значений параметров
Базы знаний (БЗ)	Базы данных (БД)
Логические зависимости между переменными	Функциональные и корреляционные зависимости между переменными
Дискретные табличные распределения	Нормальное и аналитические распределения
Алгоритмические итеративные методы решения обратных и оптимизационных задач	Аналитические методы решения обратных и оптимизационных задач
Л-функции любой сложности с любым числом объектов в статистике, параметров и градаций параметров	Функции ограниченной сложности с небольшим числом объектов, параметров и градаций параметров
Статические и динамические модели риска	Статические модели риска
Оценки по статистическим данным	Экспертные оценки
Комбинированные Л-модели риска	Отдельные модели риска
Управление по вкладам инициирующих событий в риск и эффективность	Управление по значениям риска и эффективности
Логика и дискретная математика	Теория статистики и непрерывная математика

получить систему Л-уравнений или базу знаний (БЗ), использовать ЛВ-исчисление и решать задачу риска, эффективности и управления.

1. ПЕРЕХОД ОТ БАЗЫ ДАННЫХ К БАЗЕ ЗНАНИЙ

Общим при решении перечисленных во Введении проблем является одинаковое табличное представление статистических данных (базы данных) [2, 3].

База данных (БД) табличного типа (табл. 2) содержит статистическую информацию об однородных объектах (кредитах) или состояниях одного объекта в разные моменты времени (портфель ценных бумаг). В таблице БД число столбцов может достигать нескольких десятков, а число строк — нескольких сотен. В БД значения параметра могут иметь бесконечное множество рациональных значений.

В ячейках таблицы — значения параметров объекта. Для их измерения используются шкалы: логическая, качественная, числовая и др. Последний столбец таблицы — параметр эффективности. Параметры объекта обозначим строчными буквами $z_1, \dots, z_j, \dots, z_n$, а параметр эффективности — строчной буквой $y_i, i = 1, 2, \dots, N$. В клетках табл. 2 — значения параметров z_{ji} , и в последнем столбце — параметра y_i .

База знаний. Поскольку «статистика знает все», то она и есть само знание. Нужно только умело ставить вопросы и избавиться от проклятия «бесконечно большого числа значений», которое плохо воспринимается и мало пригодно для вычислений.

Изменим исходное представление БД, заменив значения параметров на их градации (нумерованные интервалы) (табл. 3). В сценариях и ЛВ-моделях

Таблица 2

База данных — объекты и значения параметров

Объекты	Параметр 1 z_1	...	Параметр j z_j	...	Параметр n z_n	Параметр эффективности, y
1	z_{11}	...	z_{j1}	...	z_{n1}	y_1
...
i	z_{1i}	...	z_{ji}	...	z_{ni}	y_i
...
N	z_{1N}	...	z_{jN}	...	z_{nN}	y_N

База знаний — объекты и градации параметров

Объекты	Параметр 1 Z_1	...	Параметр j Z_j	...	Параметр n Z_n	Параметр эффективности Y
1	Z_{11}	...	Z_{j1}	...	Z_{n1}	Y_1
...
r	Z_{1r}	...	Z_{jr}	Y_r
...
N	Z_{1N}	...	Z_{jN}	...	Z_{nN}	Y_N

лях риска при решении проблем классификации, инвестирования, эффективности, менеджмента, взяток и управления имеется большое число N объектов (до 1000 и более), событий-параметров n (до 20 и более) и событий-градаций в каждом событии-параметре (от 2 до 40).

Таким образом, в табл. 3 для каждого параметра введены множества с конечным числом элементов (градаций). Если в табл. 2 параметры могли принимать несчетное множество разных значений, то теперь параметр имеет конечное число элементов.

Теперь параметры объекта будем называть событиями-параметрами и логическими переменными и обозначать прописными буквами $Z_1, \dots, Z_j, \dots, Z_n$, а параметр эффективности объекта — событием-параметром эффективности и обозначать буквой Y . В клетках табл. 3 находятся события-градации $Z_{jr}, j = 1, 2, \dots, n; r = 1, 2, \dots, N_j$ параметров $Z_1, \dots, Z_j, \dots, Z_n$. В последнем столбце — события-градации $Y_r, r = 1, 2, \dots, N_y$ параметра эффективности Y .

Для проблемы классификации получим систему Л-уравнений риска:

$$\begin{aligned}
 Z_{1r_1}^1 \vee \dots \vee Z_{jr_j}^1 \vee \dots \vee Z_{nr_n}^1 &= Y_{r_y}^1, \\
 &\dots \\
 Z_{1r_1}^i \vee \dots \vee Z_{jr_j}^i \vee \dots \vee Z_{nr_n}^i &= Y_{r_y}^i, \quad (1) \\
 &\dots \\
 Z_{1r_1}^N \vee \dots \vee Z_{jr_j}^N \vee \dots \vee Z_{nr_n}^N &= Y_{r_y}^N,
 \end{aligned}$$

где $i = 1, 2, \dots, N; j = 1, 2, \dots, n; r_j \in N_j; r_y \in N_y$.

Систему Л-уравнений типа (1) и будем называть базой знаний, а также рассматривать как систему Л-высказываний и использовать для получения новых знаний. С каждой Л-переменной свяжем вероятность ее истинности. Обозначим для события-параметра: p — вероятность неуспеха; $q = 1 - p$ — вероятность успеха.

Систему Л-уравнений (1) после ее ортогонализации можно преобразовать в следующую систему вероятностных полиномов:

$$\begin{aligned}
 &P_{1r_1}^1 + P_{2r_2}^1(1 - P_{1r_1}^1) + \\
 &+ P_{3r_3}^1(1 - P_{1r_1}^1)(1 - P_{2r_2}^1) + \dots = P_{r_y}^1, \\
 &\dots \\
 &P_{1r_1}^i + P_{2r_2}^i(1 - P_{1r_1}^i) + \\
 &+ P_{3r_3}^i(1 - P_{1r_1}^i)(1 - P_{2r_2}^i) + \dots = P_{r_y}^i, \quad (2) \\
 &\dots \\
 &P_{1r_1}^N + P_{2r_2}^N(1 - P_{1r_1}^N) + \\
 &+ P_{3r_3}^N(1 - P_{1r_1}^N)(1 - P_{2r_2}^N) + \dots = P_{r_y}^N.
 \end{aligned}$$

В ЛВ-теории риска с ГНС события-параметры связаны Л-операциями AND, OR, NOT и имеются циклы. В общем случае получим другие записи уравнений (1) и (2). Событиям-параметрам соответствуют Л-переменные, которые могут быть зависимыми потому, что они содержатся в Л-формуле, которая и определяет зависимость между ними. События-градации для каждого параметра являются зависимыми и образуют ГНС [2—4].

2. ВЕРОЯТНОСТИ В ГРУППАХ НЕСОВМЕСТИМЫХ СОБЫТИЙ

Запишем для объекта таблицы в общем виде Л-функцию риска и В-функцию риска

$$\begin{aligned}
 Y &= Y(Z_1, Z_2, \dots, Z_n); \\
 P_i\{Y_i = 0 | Z(i)\} &= P(P_1, \dots, P_j, \dots, P_n), \quad i = 1, 2, \dots, N,
 \end{aligned}$$

где $Y = 0$ — неуспех, $Y = 1$ — успех.

Для события-градации в ГНС рассматривают три вероятности (рис. 1): P_{2jr} — частотная вероятность градации r в j -ГНС по статистике; P_{1jr} — нормированная вероятность неуспеха объекта от градации r (сумма вероятностей в ГНС равна 1); P_{jr} — вероятность неуспеха от градации r , подставляемая в В-функцию риска вместо P_j для собы-



тия-параметра j . Определим эти вероятности для j -й ГНС:

$$P2_{jr} = P\{Z_{jr} = r\}; \quad \sum_{r=1}^{N_j} P2_{jr} = 1; \quad r = 1, 2, \dots, N_j;$$

$$P1_{jr} = P\{Y = 0 | Z_{jr} = r\}; \quad \sum_{r=1}^{N_j} P1_{jr} = 1; \quad (3)$$

$$r = 1, 2, \dots, N_j;$$

$$P_{jr} = P\{Y = 0 | Z_{jr} = r\}; \quad r = 1, 2, \dots, N_j, \quad (4)$$

где n — число признаков, N_j — число градаций в j -признаке (или обозначение множества).

Средние значения вероятностей $P2_{jr}$, $P1_{jr}$ и P_{jr} для градаций в ГНС:

$$P2_{jm} = 1/N_j; \quad P_{jm} = \sum_{r=1}^{N_j} P_{jr} P2_{jr};$$

$$P1_{jm} = \sum_{r=1}^{N_j} P1_{jr} P2_{jr}. \quad (5)$$

Вероятности P_{jr} будем оценивать при алгоритмической итеративной идентификации В-модели риска по статистическим данным. Вначале нужно определить вероятности $P1_{jr}$, удовлетворяющие выражению (3), и перейти от вероятностей $P1_{jr}$ к вероятностям P_{jr} . Число независимых вероятностей

$$N_{ind} = \sum_{j=1}^n N_j - n.$$

Вероятности P_{jr} , $P1_{jr}$, $P1_{jm}$ и P_{jm} , вычисляемые по соотношениям (3)–(5), связаны по формуле Байеса для случая ограниченного количества информации [2–4]:

$$P_{jr} = \frac{P1_{jr} P_{jm}}{P1_{jm}}; \quad r = 1, 2, \dots, N_j; \quad j = 1, 2, \dots, n.$$

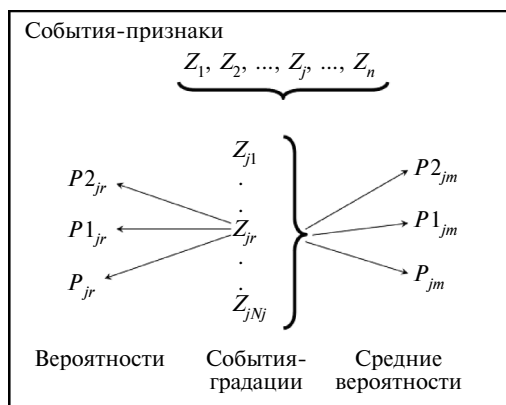


Рис. 1. Вероятности в группе несовместных событий

3. ЛОГИЧЕСКИЕ И ВЕРОЯТНОСТНЫЕ УРАВНЕНИЯ БАЗЫ ЗНАНИЙ

Объект-событие в табл. 3, описываемый Л-переменными Z_1, Z_2, \dots, Z_n , определяется следующей Л-функцией (конъюнкцией) [2, 3]:

$$Y_1 = Z_1 \wedge Z_2 \wedge \dots \wedge Z_j \wedge \dots \wedge Z_n. \quad (6)$$

Неуспех объекта-события определяется следующей Л-функцией:

$$Y_2 = Z_1 \vee Z_2 \vee Z_3 \vee \dots \vee Z_n. \quad (7)$$

Вместо Л-переменных Z_1, Z_2, \dots, Z_n в Л-выражения (6) и (7) следует подставить Л-переменные для градаций этих переменных. Из Л-функций (6) и (7) получают В-полиномы. В-функция (полином) для Л-функции (6):

$$P\{Y_1 = 1\} = q_1 q_2 q_3 \cdot \dots \cdot q_n.$$

Для перехода от Л-функции (7) к В-функции нужно в общем случае (когда Л-функция представляет собой сложное выражение) выполнить ее ортогонализацию:

$$Y_2 = Z_1 \vee Z_2 \bar{Z}_1 \vee Z_3 \bar{Z}_1 \bar{Z}_2 \vee \dots$$

Тогда В-полином запишется в виде

$$P\{Y = 0\} = p_1 + p_2(1 - p_1) + p_3(1 - p_2)(1 - p_1) + \dots \quad (8)$$

Приведенные Л-функции просты (последовательное соединение для успеха и параллельное для неуспеха), и вместо вида (8) часто используют другое выражение для В-полинома:

$$P\{Y = 1\} = 1 - q_1 q_2 \times \dots \times q_n. \quad (9)$$

Арифметика В-модели риска такова, что риск по формулам (8) и (9) находится в интервале $\{0, 1\}$ при любых значениях вероятностей инициирующих событий.

В соответствии с табл. 3 строятся системы (1) и (2) из N уравнений. Наибольшее число возможных разных объектов (логических уравнений) равно

$$N_{max} = N_1 N_2 \times \dots \times N_j \times \dots \times N_n,$$

где $N_1, \dots, N_j, \dots, N_n$ — число градаций в событиях-параметрах. Если число параметров $n = 20$ и каждый из них имеет $N_j = 5$ градаций, то число возможных разных объектов равно астрономическому числу $N_{max} = 5^{20}$, что объясняет вычислительные трудности задач риска.

Таблица 4

Типы ЛВ-моделей риска с ГНС

Типы ЛВ-моделей риска	Содержание
С полным числом событий	Л-функция риска в совершенной дизъюнктивной нормальной форме
Сограниченным числом событий	Л-функция риска в виде: кратчайших путей успеха, минимальных сечений неуспеха и ограниченного (сценарием) числа событий
Динамические	С изменяющимися вероятностями событий и временем — параметром (в техническом анализе)
Комбинированные	С логическим объединением связями OR, AND и NOT отдельных сценариев и Л-функций риска

Запишем Л-функцию для всех объектов в табл. 3

$$Y = Y_1 \vee Y_2 \vee \dots \vee Y_j \vee \dots \vee Y_N, \quad (10)$$

где каждый объект определяется Л-функцией (6) со всеми Л-переменными. В табл. 3 каждая Л-переменная в выражении (6) принимает разные значения по числу градаций, на которые она разбита. Л-функции для двух разных объектов, например,

$$Y_i = Z_1 \wedge Z_2 \wedge \dots \wedge Z_{jr} \wedge \dots \wedge Z_n,$$

$$Y_k = Z_1 \wedge Z_2 \wedge \dots \wedge Z_{jr+1} \wedge \dots \wedge Z_n,$$

ортогональны $Y_i \wedge Y_k = 0$, так как $Z_{jr} \wedge Z_{jr+1} = 0$, ибо Z_{jr} и Z_{jr+1} принадлежат одной ГНС.

Свойство ортогональности Л-слагаемых Л-функции риска (10) позволяет перейти от Л-функций к алгебраическим выражениям для вероятностей, т. е. Л-переменные заменить на вероятности и знаки « \vee » на знаки «+».

Исходя из изложенного и обобщений по разным приложениям [2–7], используются следующие типы ЛВ-моделей риска с ГНС (табл. 4).

5. ИДЕНТИФИКАЦИЯ ЛВ-МОДЕЛИ РИСКА С ГРУППАМИ НЕСОВМЕСТНЫХ СОБЫТИЙ ПО СТАТИСТИЧЕСКИМ ДАННЫМ

Задача идентификации В-модели риска по статистическим данным — обратная оптимизационная задача и является основной и сложной в исследованиях проблем риска [2, 3]. Для ее решения систему логических уравнений (1) заменяют на систему алгебраических уравнений (В-полиномов) (2). Вероятности в левой части системы уравнений

неизвестны. Задачу идентификации решают алгоритмическими итеративными методами.

Предложена следующая схема решения задачи. Пусть известны в первом приближении вероятности для градаций P_{jr} , $i = 1, 2, \dots, N_j$; $j = 1, 2, \dots, n$ и вычислены риски P_i , $i = 1, \dots, N$, для объектов в статистике из N_g «хороших» и N_b «плохих» объектов. Определим допустимый риск P_{ad} так (рис. 2), чтобы принятое расчетное число «хороших» объектов N_{gc} имело риск меньше P_{ad} и число «плохих» объектов $N_{bc} = N - N_{gc}$ имело риск больше P_{ad} . На шаге оптимизации изменим вероятности P_{jr} , $r = 1, 2, \dots, N_j$; $j = 1, 2, \dots, n$, так, чтобы число корректно распознаваемых объектов увеличилось. Переменные P_{ad} и N_{gc} связаны однозначно. В алгоритме удобнее задавать N_{gc} и определять P_{ad} .

Условие $P_i \geq P_{ad}$ выделяет следующие типы объектов: N_{gg} — «хорошие» по модели и статистике; N_{gb} — «хорошие» по модели и «плохие» по статистике; N_{bg} — «плохие» по модели и «хорошие» по статистике; N_{bb} — «плохие» по модели и статистике. Риски объектов N_{gg} , N_{gb} , N_{bg} и N_{bb} перемещаются относительно P_{ad} при изменении P_{jr} . При переходе одних объектов вправо от P_{ad} (см. рис. 2) такое же число объектов переходит влево. Оптимально такое изменение P_{jr} , которое переводит объекты N_{gb} и N_{bg} через точку P_{ad} навстречу друг другу.

Задача идентификации В-модели риска сформулирована следующим образом [2, 3].

Заданы: статистика по объектам, имеющая N_g «хороших» и N_b «плохих» объектов, и В-модель риска типа (8).

Требуется определить: вероятности P_{jr} , $j = 1, 2, \dots, n$; $r = 1, 2, \dots, N_j$, событий-градаций и допустимый риск P_{ad} , разделяющий объекты на «хорошие» и «плохие».

Критерий оптимизации: максимизация числа корректно классифицируемых объектов:

$$F = N_{bb} + N_{gg} \rightarrow \max_{P_{jr}}. \quad (11)$$

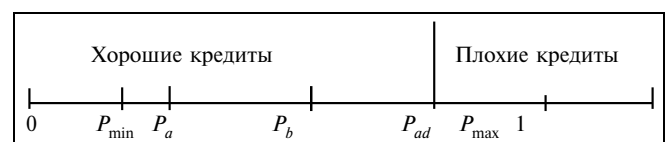


Рис. 2. Схема классификации объектов



Важно, что, в отличие от скоринговых методик, используется прямой прозрачный целочисленный критерий оптимизации.

Из выражения (11) следует, что точность В-модели риска в классификации «хороших» E_g и «плохих» объектов E_b и в целом E_m определяется как

$$E_g = N_{gb}/N_g; E_b = N_{bg}/N_b; E_m = (N - F)/N. \quad (12)$$

Ограничения:

— вероятности P_{jr} должны удовлетворять условию

$$0 < P_{jr} < 1, \quad j = 1, 2, \dots, n; \quad r = 1, 2, \dots, N_j;$$

— средние риски объектов по В-модели риска и статистике должны быть равны; при обучении В-модели риска будем корректировать вероятности P_{jr} на шаге итеративного процесса идентификации по формуле

$$P_{jr} = P_{jr}(P_{av}/P_m); \quad j = 1, 2, \dots, n; \quad r = 1, 2, \dots, N_j,$$

где $P_{av} = N_b/N$ — средний риск по статистике, P_m — средний риск по модели;

— допустимый риск P_{ad} определяется при заданном коэффициенте асимметрии распознавания «хороших» и «плохих» объектов, равном

$$E_{gb} = N_{gb}/N_{bg};$$

потери неэквивалентны при ошибочном распознавании «хорошего» и «плохого» объектов.

Формула для итеративной алгоритмической идентификации ЛВ-модели риска

$$\Delta P1_{jr} = K_1 \frac{N_{opt} - N_v}{N_{opt}} K_3 P1_{jr}, \quad (13)$$

$$j = 1, 2, \dots, n; \quad r = 1, 2, \dots, N_j,$$

где $K_1 \sim 0,05$; N_{opt} и N_v — число оптимизаций и номер текущей оптимизации, K_3 — случайное число в интервале $[-1, +1]$. При итеративной алгоритмической оптимизации величина $\Delta P1_{jr}$ стремится к нулю. Формула (13) обеспечивает простое задание максимальных приращений вероятностей, сходимость оптимизации и определение точности оценки вероятностей по их приращениям на последней оптимизации. Риск нового объекта оценивают на В-модели риска при известных вероятностях P_{jr} .

Алгоритмический итеративный метод идентификации ЛВ-модели риска с ГНС позволяет извлекать из БЗ или системы Л-уравнений новые знания (вероятности и вклады событий-градаций, допустимый риск, асимметрию распознавания и др.) при любой сложности ЛВ-функции риска и большом числе объектов в статистике.

Идентификация ЛВ-модели риска с ГНС по статистическим данным отличается высокой вычислительной сложностью и необходимы специально разработанные программные средства [2, 3].

5. АНАЛИЗ РИСКА

Пусть В-модель риска идентифицирована по статистическим данным и известны вероятности P_{jr} . Определим вклады событий-параметров и событий-градаций в риск объекта, средний риск множества объектов и в точность ЛВ-модели риска. Для этого вычислим разности между значениями названных характеристик после идентификации и при условии придания вероятностям событий-признаков нулевых значений [2, 3].

Вклад признака (всех градаций признака) в риск объекта i

$$\Delta P_j = P(i) - P(i)|_{P_j=0}, \quad j = 1, 2, \dots, n. \quad (14)$$

Вклад признака в средний риск P_m множества объектов

$$\Delta P_{jm} = P_{jm} - P_{jm}|_{P_j=0}, \quad j = 1, 2, \dots, n. \quad (15)$$

Вклад признака в целевую функцию F_{max}

$$\Delta F_j = F_{max} - F|_{P_j=0}, \quad j = 1, 2, \dots, n. \quad (16)$$

Вычисление вкладов градаций ΔP_{jrm} и ΔF_{jr} по предложенной схеме было бы некорректно, ибо не ясно, как корректировать частоты других градаций $P2_{jr}$ в ГНС, если одной из них придается нулевое значение. Поэтому вместо вкладов ΔF_{jr} будем вычислять ошибки классификации объектов по каждому событию-градации:

$$E_{jrg} = (N_{jrg} - N_{jrgg})/N_{jrg}; \quad E_{jrb} = (N_{jrb} - N_{jrbb})/N_{jrb}; \\ E_{jrm} = (N_{jr} - N_{jrgg} - N_{jrbb})/N_{jr}, \quad (17)$$

где N_{jrg} , N_{jrb} и N_{jr} — соответственно числа «хороших», «плохих» и всех объектов с градацией; N_{jrgg} и N_{jrbb} — число «хороших» и число «плохих» объектов с корректной классификацией.

6. УПРАВЛЕНИЕ РИСКОМ

Из системы Л-уравнений и соответствующей системы В-полиномов получают методом алгоритмической итеративной идентификации новые знания: вероятности P_{jr} , $j = 1, 2, \dots, n$; $r = 1, 2, \dots, N_j$; вклады событий-градаций Z_{jr} , $j = 1, 2, \dots, n$; $r = 1, 2, \dots, N_j$ в риск каждого объекта, всего мно-

жества объектов и точность ЛВ-модели риска (14)–(17); допустимый риск P_{ad} .

Атрибуты риска события-градации: вероятность неуспеха для объекта, относительная вероятность неуспеха среди градаций события-параметра, вероятность-частота в множестве объектов, вклад в точность модели.

Атрибуты риска события-параметра: вероятность неуспеха для объекта, структурный вес и значимость в модели риска, вклад в средний риск множества объектов.

Атрибуты риска объекта: риск неуспеха, возможные потери, цена за риск, вклад в риск множества объектов.

Атрибуты риска множества объектов: допустимый риск, средний риск, средние потери, допустимые потери, асимметрия распознавания хороших и плохих объектов, число объектов, число опасных объектов, энтропия рисков.

В ЛВ-теории риска с ГНС используется активное управление риском и эффективностью, исходя из атрибутов риска объектов, объекта, параметров объекта и градаций параметров. Управлять — значит знать, что и насколько изменить для достижения допустимого риска или эффективности. Управляют принятием очевидного решения.

Кредитным риском управляют, исходя из количественных значений описанных атрибутов. Цель управления состоит в снижении финансовых потерь банка и повышении точности распознавания плохих и хороших кредитов. Параметры управления риском кредита и кредитной деятельностью банка — риск кредита, который сравнивают с допустимым риском и принимают решение о выдаче кредита; коэффициент асимметрии распознавания хороших и плохих кредитов; цена за кредит, зависящая от риска кредита и допустимого риска; число параметров, описывающих кредит; число градаций для каждого параметра; ширина интервалов при выделении градаций для таких параметров, как сумма и срок кредита, возраст клиента и др.

В работах [2–4] приведены результаты идентификации ЛВ-модели кредитных рисков по реальным статистическим данным банка. Имелось $N = 1000$ кредитов, из которых 700 было «хороших» и 300 «плохих». Оценивались 94 вероятности P_{jr} и допустимый риск P_{ad} . ЛВ-модель кредитного риска с ГНС показала в два раза большую точность и в семь раз большую робастность, чем известные методики, а также абсолютную прозрачность в распознавании плохих и хороших кредитов.

Логико-вероятностную модель кредитного риска используют для оценки риска новых кредитов по выражению (8), в которое для конкретного кредита подставляют вероятности событий-градаций

параметров. Вклады в риск кредита пропорциональны значениям этих вероятностей. По значению риска кредита устанавливают, «хороший» это или «плохой» кредит, какие градации и насколько влияют на это. Такую модель используют также для управления кредитной деятельностью банка. Приведем примеры.

Пример 1. Выполним оценку одного кредита по данным анкеты на его получение с помощью программного комплекса для лабораторных работ. Кредиты описываются следующими параметрами: Z_1 — сумма счета в банке (4 градации), Z_2 — срок займа (10 гр.), Z_3 — кредитная история о возвращении займов (5 гр.), Z_4 — назначение займа (11 гр.), Z_5 — сумма займа (10 гр.), Z_6 — счета по ценным бумагам (5 гр.), Z_7 — продолжительность работы (5 гр.), Z_8 — взнос в частичное погашение (4 гр.), Z_9 — семейное положение и пол (4 гр.), Z_{10} — совместные обязательства или поручитель (3 гр.), Z_{11} — время проживания в данной местности (4 гр.), Z_{12} — вид гарантии (4 гр.), Z_{13} — возраст (5 гр.), Z_{14} — наличие других займов (3 гр.), Z_{15} — наличие жилой площади (3 гр.), Z_{16} — число зай-

Таблица 5

Параметры кредита

Параметр	Градация	Вероятность P_{jr}
1	4	0,004997
2	3	0,036728
3	2	0,035466
4	7	0,047464
5	7	0,012243
6	4	0,002746
7	3	0,019102
8	2	0,012360
9	2	0,016730
10	1	0,038348
11	2	0,021220
12	2	0,020309
13	2	0,020840
14	1	0,020590
15	1	0,024678
16	3	0,017928
17	4	0,021060
18	1	0,017080
19	2	0,016327
20	1	0,018892



мов с банком, включая новый заем (4 гр.), Z_{17} — профессия (4 гр.), Z_{18} — число родственников на иждивении (2 гр.), Z_{19} — наличие телефона (2 гр.), Z_{20} — иностранный или местный житель (2 гр.).

Конкретный кредит имел параметры, представленные в табл. 5. Вероятности P_{jr} получены решением задачи идентификации по статистическим данным банка. Кредит получил следующие оценки: риск равен 0,350070 и больше допустимого риска $P_{ad} = 0,306274$; кредит плохой, но его можно дать под 18 % (вместо 10 % при P_{ad}). Риск кредита получен по формуле (8). Вклады признаков и градаций в риск кредита пропорциональны значениям P_{jr} .

Пример 2. *Оптимальное число параметров для описания кредита* [2]. Для каждого параметра j после обучения модели риска определялись: средние значения вероятностей P_{1jm} и P_{jm} (5), а также уменьшение числа корректно распознанных хороших и плохих кредитов при исключении этого параметра из модели риска (16). Модель после этого переобучалась. Наибольший вклад в точность распознавания кредитов вносят события-параметры $Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_{13}$. Нулевой вклад вносят события-параметры $Z_{11}, Z_{12}, Z_{17}, Z_{18}, Z_{19}$. Их можно исключить из описания кредита.

Пример 3. *Оптимальное число событий-градаций N_j в параметре.* Зависимость точности ЛВ-модели риска от числа градаций в параметре исследовалась для срока кредита, который в исходном варианте имеет 10 градаций. После переобучения модели риска получены следующие результаты: при отсутствии параметра $F_{\max} = 800$; при двух градациях в параметре $F_{\max} = 812$; при четырех градациях — $F_{\max} = 816$; при десяти градациях — $F_{\max} = 824$; при ста градациях — $F_{\max} = 828$. Для каждого параметра кредита для каждого банка имеется оптимальное число градаций.

7. ПРИЛОЖЕНИЯ ЛВ-МОДЕЛЕЙ РИСКА С ГРУППАМИ НЕСОВМЕСТИМЫХ СОБЫТИЙ

Объектами исследования были структурно-сложные, многокомпонентные и многоуровневые системы: технические, банковские, экономические и организационные [5]. Дадим краткое описание приложений ЛВ-моделей риска с ГНС: кредитные риски, инвестирование, эффективность, менеджмент, взятки и управление развитием [2–4].

Кредитные риски. Использовались статистические данные одного западного и двух российских банков [2–4]. В статистиках было от 500 до

1000 кредитов. Логико-вероятностные модели риска показали в два раза большую точность в распознавании хороших и плохих кредитов; в семь раз большую устойчивость в классификации кредитов, абсолютную прозрачность в оценке и анализе риска и модели риска. Для западного банка кредитный риск уменьшался с 28 до 17 %. Для российских банков кредитный риск уменьшался примерно вдвое.

Инвестирование. Разработана ЛВ-модель риска инвестиций в портфель ценных бумаг. В этой модели, в отличие от теорий Марковица и Value-at-Risk (VaR) (стоимость под риском), не используется допущение о нормальности законов распределения доходности ценных бумаг и всего портфеля. Используется переход от табличной БД по доходностям акций портфеля к БЗ в виде системы Л-уравнений. Для этого диапазон изменения доходности каждой акции и портфеля разбивается на интервалы (события-градации). Наряду с допустимой доходностью портфеля Y_{ad} и риском портфеля $Risk$, вводятся понятия числа опасных состояний N_{ad} и энтропии H_{ad} рисков опасных состояний в «хвосте» распределения доходности портфеля. Это позволяет анализировать риск и эффективность (доходность) портфеля и управлять ими [2, 3].

Эффективность. Разработана ЛВ-модель риска эффективности социальных процессов, в которой оцениваются весомости инициирующих случайных процессов для итогового случайного процесса параметра эффективности. Используется переход от статистической БД значений влияющих параметров и итогового параметра эффективности к БЗ в виде системы Л-уравнений. Для этого диапазоны изменения влияющих и итогового параметров разбиваются на интервалы (события-градации). Для параметра эффективности вводятся понятия допустимого значения Y_{ad} , риска $Risk$, числа опасных состояний N_{ad} и энтропии H_{ad} рисков опасных состояний в «хвосте» распределения параметра эффективности. Это позволяет анализировать риск и эффективность и управлять ими. В отличие от работ лауреата Нобелевской премии Джеймса Хекмана, задача решается при произвольных распределениях влияющих параметров и параметра эффективности [2, 3].

Менеджмент. Разработаны ЛВ-модели риска неуспеха менеджмента компании по функциям, по направлениям деятельности, по управлению компанией как сложным объектом, по достижению одной и группы целей, по качеству функционирования. В рассмотренных моделях статистические данные не используются и не осуществляется переход от БД к БЗ. Демонстрируются построение и

Особенности ЛВ-теории риска с ГНС

Особенности	Содержание
Объекты исследования	Структурно-сложные, многокомпонентные и многоуровневые системы: банковские, экономические и организационные
ЛВ-модели риска с ГНС	С полным числом событий, с ограниченным числом событий, комплексные модели риска, динамические модели риска
Области применения	Проблемы классификации, инвестирования, эффективности, менеджмента, взяток и коррупции, управления развитием
Решаемые задачи	Количественная оценка и анализ риска объекта и множества объектов, управление риском, управление риском и доходностью портфеля, моделирование риска неуспеха менеджмента, выявление взяток
Методические основы ЛВ-теории риска с ГНС	Логика, теории множеств и вероятностей, комбинаторика, дискретная математика, ЛВ-исчисление с логическими связями элементов OR, AND, NOT и циклами
Переход от базы данных к базе знаний	Вводят ГНС или конечные множества и преобразуют статистическую БД в БЗ в виде системы Л-уравнений, что позволяет использовать ЛВ-исчисление и решать задачи риска, эффективности и управления
Определение риска	Вводится допустимый риск, разделяющий объекты на «плохие» и «хорошие». Риск каждого уровня определяется своими атрибутами
Критерии качества Л-модели риска	Точность (ошибки классификации), робастность, прозрачность результатов и анализа риска и модели риска
Распределения переменных	Дискретные табличные распределения случайных переменных
Связанность и зависимость переменных	Л-переменные зависимы, но не изначально, а только потому, что они содержатся в определенной Л-формуле, определяющей зависимость между ними. События-градации для каждого параметра зависимы и рассматриваются как ГНС
Тип вычислений	Алгоритмические интерактивные методы. Обратная оптимизационная задача решается независимо от числа объектов в БД, числа параметров в объекте, и числа градаций в каждом параметре, Л-сложности модели риска
Построение модели риска	Сценарий, граф-модель риска, Л-функция риска, ортогональная Л-функция, В-функция риска, идентификация В-функции риска
Идентификация ЛВ-модели риска	Идентификация В-модели риска по статистическим данным выполняется путем решения задачи оптимизации алгоритмическими итеративными методами, использующими формулу Байеса
Логические разности	На каждом шаге оптимизации нужно определить приращение каждой вероятности. При этом следует нормировать вероятности в ГНС, т. е. их сумма должна равняться 1
Анализ риска	Вычисляют вклады в риск объекта, множества объектов и целевую функцию событий-параметров и событий-градаций. Эти вклады — атрибуты риска — и позволяют управлять риском и эффективностью
Управление риском	Активное управление риском по вкладам событий-признаков и событий-градаций в риск и эффективность
Программные средства	Специальные программные обеспечения для идентификации В-модели риска, выбора портфеля ценных бумаг, структурно-логического моделирования



использование комплексных ЛВ-моделей риска, построенных из частных моделей. На необходимость разработки моделей риска для управления менеджментом указывал в своих работах выдающийся американский экономист Питер Друкер [2].

Взятки. Разработаны следующие ЛВ-модели риска взяток: учреждения по параметрам его функционирования; чиновника по параметрам его поведения; учреждения и чиновника по параметрам обслуживания; Л-сложения первых перечисленных моделей. Приведены примеры этих ЛВ-моделей риска взяток, в том числе по выявлению риска взяток по реальным данным приема в детский садик. Параметром обслуживания было принято время ожидания ребенка в очереди. Для построения названных ЛВ-моделей риска применяется переход от табличных статистических БД к БЗ в виде системы Л-уравнений. Данные модели предназначены для департамента «Экономические преступления» города, служб внутренней безопасности компаний и банков и для разработки стандартов на параметры обслуживания [6].

Управление развитием организационных и экономических систем, сложных машин и технологий осуществляется по критериям риска и эффективности. Процесс развития рассматривается как процесс управления сложным объектом с движением к заданному конечному состоянию из начального по выбранной траектории и с коррекцией в случае отклонения от нее. Приведены примеры построения сценариев и моделей риска, программ развития. Анализ процессов доводочных испытаний многих машин и систем показал, что КПД процесса доводки (отношение времени испытаний к календарному времени) равен 5—12 %. Предложенная технология управления процессом развития существенно повышает его эффективность [7].

ЗАКЛЮЧЕНИЕ

На основе изложенного и результатов работ [2—7] в табл. 6 в качестве обобщения сформулированы особенности ЛВ-теории риска с ГНС. Именно эти особенности определяют достоинства и области приложений.

Отметим также, что исследование проблем риска в экономике, технике и экологии, государственной и национальной безопасности невозможно без построения сценариев, логических и вероятност-

ных моделей риска, количественного моделирования и анализа риска и эффективности, обеспечивающих возможность управления.

Развитие работ и исследований по ЛВ-теории риска для целей управления риском и эффективностью, по нашему мнению, должно заключаться в следующем:

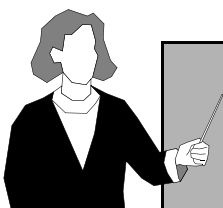
- разработка ЛВ-моделей риска с ГНС для других приложений;
- разработка учебного курса по ЛВ-теории риска с ГНС;
- совершенствование методик и алгоритмов вычислений, а также программных средств для ЛВ-моделирования и управления риском и эффективностью;
- экспертиза и сертификация ЛВ-моделей риска и программных средств.

ЛИТЕРАТУРА

1. *Рябинин И.А.* Надежность и безопасность структурно-сложных систем. — СПб.: Изд-во СПб. ун-та, 2007. — 276 с.
2. *Соложенцев Е.Д.* Сценарное логико-вероятностное управление риском в бизнесе и технике. 2-е изд. — СПб.: Бизнес-пресса, 2006. — 560 с.
3. *Solojntsev E.D.* Scenario Logic and Probabilistic Management of Risk in Business and Engineering. — Springer, 2004. — 391 p.
4. *Соложенцев Е.Д., Степанова Н.В., Карасев В.В.* Прозрачность методик оценки кредитных рисков и рейтингов. — СПб.: Изд-во СПб. ун-та, 2005. — 197 с.
5. *Соложенцев Е.Д., Махутов Н.А.* Логико-вероятностные модели риска в многокомпонентных системах с группами несовместных событий для задач классификации, инвестирования, эффективности и менеджмента // Проблемы безопасности и чрезвычайных ситуаций. — 2006. — № 3. — С. 30—52.
6. *Соложенцев Е.Д.* Сценарные логико-вероятностные модели риска взяток // Финансы и бизнес. — 2007. — № 1. — С. 125—138.
7. *Махутов Н.А., Соложенцев Е.Д.* Управление риском при испытаниях сложных машин, систем и технологий // Проблемы машиностроения и автоматизации. — 2008. — № 1.

☎ (812) 321-47-66, 📠 (812) 321-47-71,
e-mail: risk@sapr.ipme.ru, safety@imash.ru

Статья представлена к публикации членом редколлегии
Б.Г. Воликом. □



Читайте в следующем номере статью

Пономаренко В.С., Листрового С.В. Метод решения задачи о минимальном покрытии как средство планирования в GRID