

С. В. САВКОВ, В. М. ШИШКИН

## РАЗРАБОТКА СИСТЕМЫ ИНТЕРВАЛЬНОГО ОЦЕНИВАНИЯ ИНФОРМАЦИОННЫХ РИСКОВ

Представлена система интервального оценивания информационных рисков на основе разнородных и неполных исходных данных. Показана необходимость распараллеливания вычислений и описана реализация разработанной системы на вычислительном кластере.

*Ключевые слова:* безопасность, оценка рисков, гетерогенность информации, параллельные вычисления.

**Введение.** Для решения задач анализа и оценивания информационных рисков используются различные экспертные системы [1, 2], обладающие достаточно широкими функциональными возможностями и удобством в эксплуатации, что обуславливает привлекательность использования таких систем на практике [3]. В то же время они имеют существенные недостатки, характерные для многих прикладных систем экспертного оценивания: сомнительность выбора исходных данных и отсутствие характеристик рассеяния рассчитываемых показателей, что снижает достоверность оценок и доверие к результатам анализа.

В реальных условиях исходная информация плохо структурирована, неполна, неточна, часто имеет нечисловой характер, все первичные данные, по сути, являются случайными величинами. Следовательно, чтобы повысить достоверность оценок необходимо, во-первых,

предусмотреть возможность обработки именно таких, недостаточно определенных и разнородных данных и, во-вторых, обеспечить расчет стохастических характеристик показателей, используемых для принятия решений. Зачастую источники первичных данных имеют нечисловое выражение — ординальное или лингвистическое, в предельном случае это могут быть и слабоструктурированные тексты на естественном языке. В общем случае в качестве исходных данных может выступать различная информация, получаемая, например, в результате экспертизы объекта оценивания.

**Структурная метамодель.** В настоящей статье рассматривается система интервального оценивания информационных рисков. В разработанной системе в качестве основы использовалась модель, допускающая различные интерпретации [4]. Она построена на дихотомической оппозиции: „защищаемый объект“ — потенциально враждебная „среда“ в широком смысле этих слов. Подчеркивается необходимость фиксации „границы объекта“, или функционально — „границы ответственности“, и „внешней границы среды“, ограничивающей зону досягаемости для противодействия факторам риска. Элементы модели определяются в терминах трех категорий: субъектов, объектов и воздействий первых на вторые. Соответственно категориям выделяются три непересекающихся непустых подмножества множества  $M_0 = M_s \cup M_e \cup M_c$  элементов модели:

— независимые активные субъекты, „источники угроз“ — множество  $M_s$  (threat sources);

— проводники воздействий: события, порождаемые источниками угроз, „угрозы“ нарушения безопасности — множество  $M_e$  (threat events), в котором выделяется подмножество так называемых „событий риска“ — угроз, наносящих ущерб непосредственно объекту;

— компоненты объекта — множество  $M_c$  (components).

Структурная схема метамодели представлена на рис. 1. На множестве  $M_0$  определяется бинарное отношение причинности  $R$  со свойством транзитивности, к которому можно свести многие связи, имеющие имплицативный характер. Отношение  $R$  упорядочивает множество  $M_0$  и задает на нем структуру, фиксирующую каналы распространения потоков угроз от источников до объекта, и порождает квадратную матрицу отношений  $W_0$ .

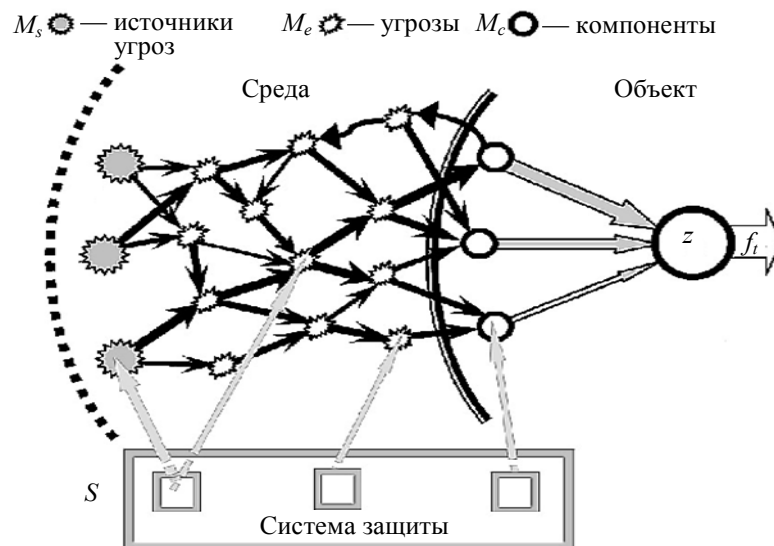


Рис. 1

Решение задачи противодействия факторам риска реализуется с помощью системы защиты информации (СЗИ), представляемой в виде множества элементов  $S$ , каждый из которых осуществляет воздействие на элементы множества  $M_0$ . Между элементами множеств  $S$

и  $M_0$  устанавливается отношение, формально сводимое к  $R$ , порождающее прямоугольную матрицу отношений  $\mathbf{R}_0$ .

Источники угроз  $M_s$  считаются генераторами потока событий (угроз), распространяющегося по каналам, заданным соотношением  $R$  на множестве  $M_0$ . Элементы  $M_e$  рассматриваются как функциональные преобразователи, перераспределяющие потоки событий. На выходе элементов  $M_e$ , представляющих события риска, формируется поток угроз, непосредственно воздействующий на объект в составе множества  $M_c$ . Тогда средства защиты  $S$  можно интерпретировать как линейные фильтры.

Роль условного элемента  $z$ , соответствующего состоянию объекта в целом как преобразователя, ограничивается функцией сумматора-интегратора. Тогда на выходе  $z$  можно фиксировать результирующий поток  $f_t$ , интеграл от которого по некоторому интервалу времени является, по сути, мерой риска для объекта, определяемой ущербом, наносимым ему за это время.

В простейшей количественной интерпретации метамоделли матрица  $\mathbf{W}_0$  отображается в арифметическую матрицу  $\mathbf{W} = (w_{ij})$ , элементы которой можно рассматривать как весовые коэффициенты, представляющие собой меру влияния  $i$ -го элемента на  $j$ -й. Матрица  $\mathbf{W}$  содержит все исходные данные для последующих расчетов.

Далее рассчитываются показатели  $v_{ij}$ , аналогичные по смыслу коэффициентам  $w_{ij}$ , но с учетом транзитивности отношений, прежде всего на состоянии  $z$ . В результате определяется матрица  $\mathbf{V}$ , структурно эквивалентная  $\mathbf{W}$ . При отсутствии рефлексии элементов, если  $\mathbf{W}$  считать взвешенной матрицей смежности некоторого графа, они легко рассчитываются как суммы по всем путям из  $i$ -й в  $j$ -ю вершину произведений оценок дуг каждого пути, что в простейшем случае равносильно матричному преобразованию  $\mathbf{V} = (\mathbf{I} - \mathbf{W})^{-1} - \mathbf{I}$ , где  $\mathbf{I}$  — единичная матрица.

Последний, всегда ненулевой,  $z$ -й столбец  $\mathbf{v}_z$  матрицы  $\mathbf{V}$  содержит искомые показатели  $\{v_{iz}\}$  влияния любого  $i$ -го фактора риска на состояние защищенности объекта. В соответствии с этими показателями выбирается СЗИ, ориентированная на противодействие наиболее значимым факторам риска.

Рассмотренный программный продукт, в котором используется подобный алгоритм, показал практическую применимость для анализа сложных объектов. Однако, обладая некоторыми преимуществами по сравнению с существующими системами анализа рисков, он имеет и недостатки, отмеченные выше.

**Алгоритм интервального оценивания.** В качестве исходной идеи для алгоритма был выбран метод арифметизации ординальных отношений, используемый в методе анализа и синтеза показателей при информационном дефиците [5] и применяемый только для простых расслоенных или древовидных структур факторов риска. Указанный способ позволяет при наличии нечисловой либо неполной информации об отношениях вычислить значения их математических ожиданий и дисперсий для последующего использования при расчете. Разработанный алгоритм позволяет арифметизировать каждый столбец матрицы  $\mathbf{W}$  по отдельности, а сама матрица  $\tilde{\mathbf{W}}$  в этом случае формируется как композиция случайных вектор-столбцов. Прямое использование этого метода для решения поставленной задачи связано с проблемой зависимости между генерируемыми элементами различных столбцов матрицы  $\tilde{\mathbf{W}}$ .

Для выполнения требования независимости при реализации рассматриваемого алгоритма будем считать, что компоненты вектора весовых коэффициентов  $\mathbf{w} = (w_1, \dots, w_m)$  имеют рав-

номерное распределение, а сам вектор  $\mathbf{w}$  также равномерно распределен на симплексе размерностью  $(m-1)$  в  $m$ -мерном пространстве. Значения параметров  $w_j$  по каждой оси отсчитываются дискретно с шагом  $h=1/n$ , где  $n$  — параметр, задающий точность представления значений. В результате генерируется множество  $W(m,n)$  всех возможных векторов весовых коэффициентов на гиперкубе и выбираются только векторы, принадлежащие симплексу.

Учет нечисловой, неточной и неполной информации  $I$  о весовых коэффициентах  $w_1, \dots, w_m$  позволяет, как правило, существенно сократить множество  $W(m,n)$  до некоторого непустого множества  $W(m,n;I)$  всех допустимых весовых векторов.

Для  $m=3$  полученное множество  $W(m,n;I)$  можно наглядно изобразить на пространственном графике, как показано на рис. 2. Условие нормировки  $w_1 + w_2 + w_3 = 1$  задает в координатах  $(w_1, w_2, w_3)$  плоскость, на которой расположены элементы множества  $W(m,n)$ . При задании различного вида исходной информации (точечной, интервальной, порядковой) формируется часть плоскости, содержащая элементы  $W(m,n;I)$ .

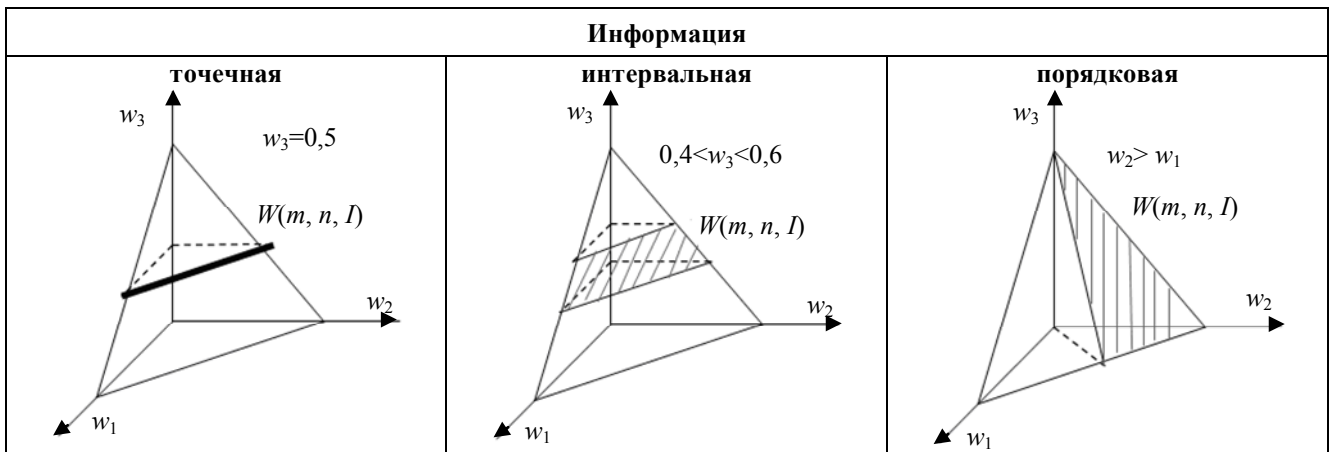


Рис. 2

Преимущество разработанного алгоритма заключается в том, что таким способом можно задать любые недостаточно определенные исходные данные, в том числе и в лингвистическом представлении. Способ задания исходной информации влияет на форму геометрической фигуры — части плоскости, содержащей множество  $W(m,n;I)$ .

Неопределенность выбора вектора  $\mathbf{w} = (w_1, \dots, w_m)$  из множества  $W(m,n;I)$  моделируется путем рандомизации этого выбора, в результате которой весовые коэффициенты преобразуются в случайные величины  $\tilde{w}_1(I), \dots, \tilde{w}_m(I)$ , имеющие совместное равномерное распределение на множестве  $W(m,n;I)$ .

**Параллельная реализация алгоритма.** Разработанная система интервального оценивания информационных рисков построена по архитектуре клиент-сервер. Серверная часть обеспечивает взаимодействие клиента с вычислительным кластером, параллельная обработка данных на кластере поддерживается средствами интерфейса MPI [6].

В разработанной программе на кластере реализована наиболее ресурсоемкая часть приложения, а именно генерация множества допустимых (удовлетворяющих исходным данным) векторов  $\mathbf{w}_j$ . Поскольку эти векторы взаимонезависимы, их вычисление может быть организовано в параллельных ветвях алгоритма, структурная схема которого представлена на рис. 3. На схеме  $i$ -я ветвь алгоритма, порожденная функцией  $\text{MPI\_Init}()$ , генерирует множество допустимых векторов для столбца  $\mathbf{w}_j$  матрицы  $\mathbf{W}$ . Каждая ветвь содержит цикл из  $N$  итераций

( $k$  — номер текущей итерации), где  $N$  определяет количество генерируемых векторов. С увеличением  $N$  возрастает точность получаемого результата. На каждой итерации генерируется случайный вектор  $\tilde{w}_j$ , имеющий равномерное распределение на симплексе в  $m$ -мерном пространстве. Далее, вектор  $\tilde{w}_j$  проверяется на соответствие исходной информации  $I$ .

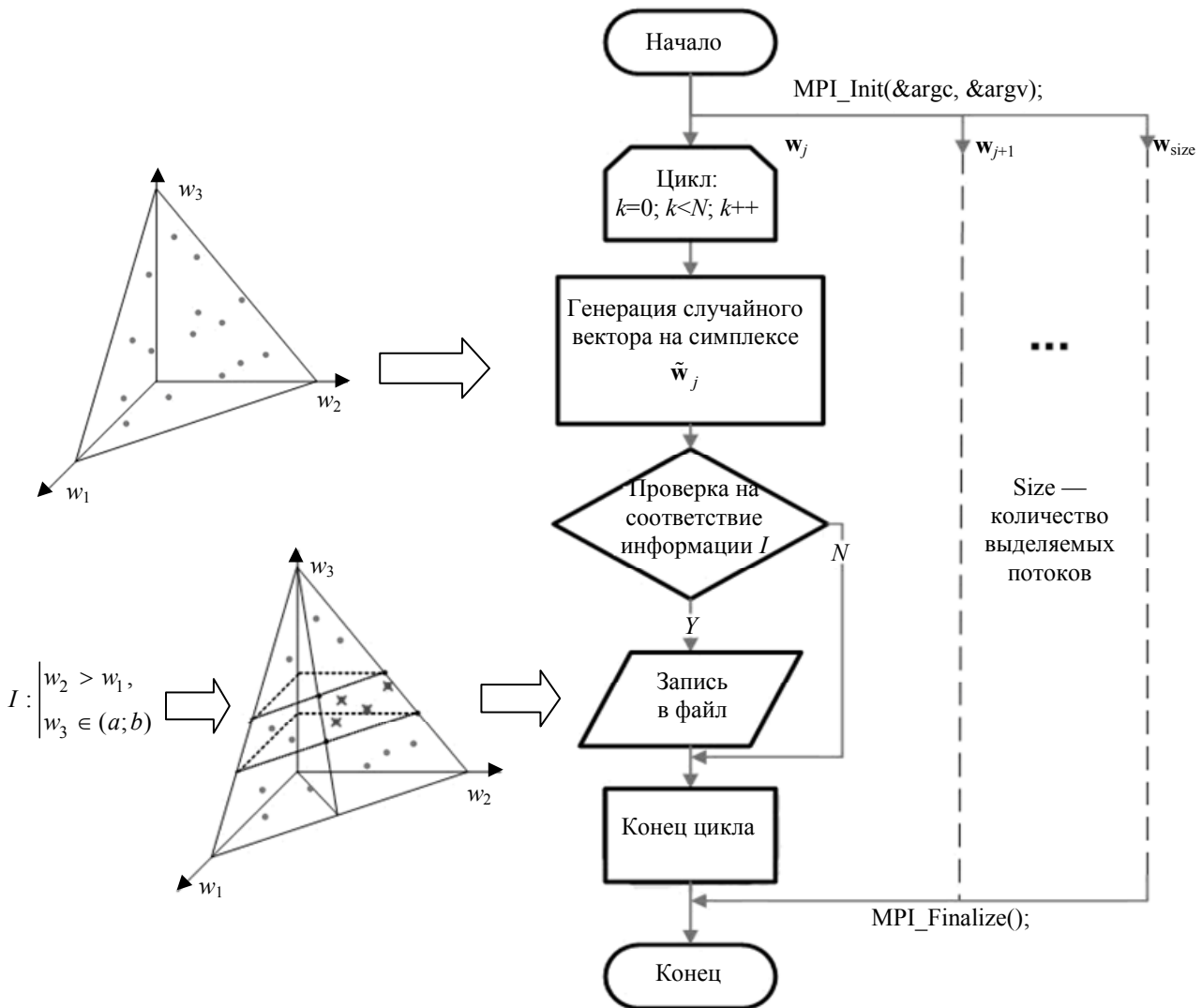


Рис. 3

Программа может работать с такими видами исходной информации, как точечная, интервальная и порядковая, допуская отсутствие информации. Расширение этого списка не представляет принципиальных сложностей. При этом если какой-либо параметр задан точно ( $w_{ij} = a$ ), то при анализе в программе это условие задается интервалом  $w_{ij} \in (a - h/2; a + h/2)$ . Векторы, удовлетворяющие информации  $I$ , записываются в файл результатов, который по окончании вычислений передается на клиентскую сторону.

**Выводы.** Разработанная система позволяет преодолеть такие недостатки существующих программных продуктов, как точечность оценок и ограничения на способы задания исходных данных. По сравнению с методиками, применяемыми в других экспертных системах, использование предложенной системы позволяет повысить достоверность оценок и доверие к результатам оценивания.

Обработка большого количества информации при анализе входных данных требует значительных вычислительных ресурсов, причем время вычислений резко возрастает с увеличе-

нием сложности объекта анализа. Параллельная реализация алгоритма позволяет за приемлемое время производить расчет более сложных моделей, представляющих реальные объекты.

Разработанная система может быть востребована организациями, проводящими комплексный аудит информационной безопасности; государственными органами, ответственными за обеспечение безопасности информационных ресурсов, комплексной и экологической безопасности объектов; вузами, ведущими подготовку по соответствующим специальностям.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Медведовский И. Д.* Современные методы и средства анализа и контроля рисков информационных систем компаний. [Электронный ресурс]: <<http://www.bugtraq.ru/library/security/itrisk.html>>.
2. *Бурдин О. А., Кононов А. А.* Комплексная экспертная система управления информационной безопасностью // Информационное общество. 2002. № 1. С. 38—44.
3. *Петренко С. А., Симонов С. В.* Управление информационными рисками // Экономически оправданная безопасность М.: ДМК Пресс, 2004. 384 с.
4. *Шишкин В. М.* Мета модель анализа, оценки и управления безопасностью информационных систем // Проблемы управления информационной безопасностью: Сб. трудов Ин-та системного анализа РАН. М.: Едиториал УРСС, 2002. С. 92—105.
5. *Хованов Н. В.* Анализ и синтез показателей при информационном дефиците. СПб: Изд-во СПбГУ, 1996. 196 с.
6. *Воеводин В. В., Воеводин Вл. В.* Параллельные вычисления. СПб: БХВ-Петербург, 2004. 606 с.

#### *Сведения об авторах*

**Сергей Витальевич Савков**

— аспирант; Санкт-Петербургский государственный университет информационных технологий, механики и оптики, кафедра проектирования компьютерных систем; E-mail: [sergsavkov@gmail.com](mailto:sergsavkov@gmail.com)

**Владимир Михайлович Шишкин**

— канд. техн. наук, доцент; Санкт-Петербургский институт информатики и автоматизации РАН, лаборатория информационно-вычислительных систем и проблем защиты информации; ст. науч. сотрудник; E-mail: [vms@iias.spb.su](mailto:vms@iias.spb.su)

Рекомендована кафедрой  
проектирования компьютерных систем  
СПбГУ ИТМО

Поступила в редакцию  
08.02.10 г.