С. В. Кудинов (Москва, ИИНиТБ, РГГУ). **Методика анализа взвешенно-факторных моделей принятий решений в области информационной безопасности.**

Взвешенно-факторные модели (ВФМ) находят широкое применение в таких областях, как экономическое моделирование [1], информационные технологии [2], информационная безопасность (ИБ) [3], управление проектами [4]. Однако в академической литературе недостаточно освещены методики их построения и анализ их свойств. В области принятия решений по проектам ИБ актуальные вопросы субъективизма и оцифровки выгод от ИБ обычно решают, используя формальные модели типа ВФМ [3]. Однако было показано [5], что, несмотря на наглядность и понятность их логики, использование этого класса систем может быть неэффективным в принятии решений по проектам ИБ. Одним из решений означенной проблемы является подход к оценке ВФМ на устойчивость и чувствительность. В работе, представленной данным сообщением, предложены алгоритмы для такого анализа.

Формализация ВФМ. Модель $M = (\mathbf{I}, \Delta)(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ — функция, преобразующая оценки объектов $\mathbf{X}$ (операционные переменные) по набору факторов в соответствии с их весами $\mathbf{Y}$ (стратегическими переменными 1-го уровня) в средневзвешенную оценку — набор индексов $\mathbf{I}$ [5]. Модель интерпретирует $\mathbf{I}$ путем его отнесения к одному из разбиений множества интерпретаций Δ. Каждое разбиение D_i ВФМ конфигурируется стратегическими переменными 2-го уровня $\mathbf{Z}$. Решение ВФМ $D_i(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ — это множество, которому принадлежит $\mathbf{I}(\mathbf{X}, \mathbf{Y})$.

Ограничимся 2 классами интерпретаций: матрицей принятия решений (МПР) и мультипликативной интерпретацией (МИ). Для МПР множество значений $\mathbf{I}$ разбивается на квадранты, образуя «сетку». В качестве $\mathbf{Z}$ используются границы разбиения множества значений отдельных компонентов $\mathbf{I}$. В МИ выделяется подмножество из $\mathbf{I}$, для которого известно, что уменьшение одного из индексов компенсируется пропорциональным ростом другого. В качестве $\mathbf{Z}$ выступает конечное множество фронтов, разделяющих множество значений $\mathbf{I}$ на гиперболические сегменты.

Пусть задана модель $M = (\mathbf{I}, \Delta) = (\mathbf{I}, D_1, D_2)(\mathbf{X}, \mathbf{Y}, \mathbf{Z}_1, \mathbf{Z}_2)$, где D_1 относится к классу МПР с границами $\mathbf{Z}_1$, а $\mathbf{D}_2$ — к МИ с фронтами $\mathbf{Z}_2$. В нашем случае параметры $\mathbf{Y}, \mathbf{Z}_1, \mathbf{Z}_2$ считаются заданными и фиксированными (частный случай анализа адекватности).

Пусть задан вектор $\mathbf{X}' = (x'_{ij})$. Тогда $\mathbf{X}'' = (x''_{ij})$ называется *одиночным отклонением* $\mathbf{X}'$ ($\mathbf{X}'' = \zeta(\mathbf{X}')$), если найдутся такие i_0, j_0, что $|x''_{i_0 j_0} - x'_{i_0 j_0}| = 1$ и для всех $i \in \{1, 2, \dots, n\}$, $j \in \{1, 2, \dots, q\}$, $i \neq i_0$, $j \neq j_0$ выполняется $|x''_{ij} - x'_{ij}| = 0$. Применение к вектору $\mathbf{X}'$ одиночного отклонения n раз обозначается $\zeta^n(\mathbf{X}')$. Будем говорить, что вектор $\mathbf{X}''$ *находится на расстоянии* h *от* $\mathbf{X}'$, если $h = \arg\min_{h \in N_0} \{\mathbf{X}'' = \zeta^h(\mathbf{X}')\}$.

Введем расстояние $|a - b|$ на множествах D_1, D_2 как разность $|\mu(a) - \mu(b)|$ по модулю мер, введенных следующим образом:

$$\mu_1(D_1(z_{1i_1}, z_{2i_2}, \dots, z_{ni_n})) = \begin{cases} 0 \text{ при } (i_1, i_2, \dots, i_n) = (1, 1, \dots, 1), \\ \max\limits_{l_t < i_t,\, t=1,\dots,n} \{D_2(z_{1l_1}, z_{2l_2}, \dots, z_{nl_n})\} + 1 \text{ иначе}, \end{cases}$$

$$\mu_2(D_2(z_i)) = \begin{cases} 0, & i = 1, \\ D_1(z_{i-1}) + 1, & i = 2, \dots, q. \end{cases}$$

Анализ устойчивости. Используем однократное колебание входных параметров и находим результирующее отклонение решения. Показатель устойчивости — эмпирическая плотность распределения отклонений на пространстве значений расстояний по колебаниям решений ВФМ.

Алгоритм. Задается число экспериментов T. Для каждого $D \in \Delta$:

— генерируется случайным образом вектор $\mathbf{X}'_t$ (T раз);

— для каждого $\mathbf{X}'_t$: вычисляется $D'_t = D(\mathbf{X}'_t)$; производится полный перебор одиночных колебаний $\mathbf{X}_t''$ от вектора $\mathbf{X}_t'$;

— для каждого $\mathbf{X}_t''$: вычисляется $D_t'' = D(\mathbf{X}_t'')$; фиксируется расстояние $|D_t' - D_t''|$; на основе статистики по $|D_t' - D_t''|$ формируется вектор эмпирических частот $\mathbf{V}_D = (v_1, v_2, \ldots)$, где v_i — отношение числа исходов, когда расстояние $|D_t' - D_t''| = i$, к числу всех исходов.

Критерий устойчивости. Модель с интерпретацией D обладает хорошей устойчивостью, если для $\mathbf{V}_D$ справедливо: 1) $v_3 + v_4 + v_5 + \cdots < 5\%$, 2) $v_2 < 15\%$.

Анализ чувствительности. Определяем, насколько минимально нужно отклонить входные операционные параметры $\mathbf{X}$ (в пределах допустимых значений), чтобы решение ВФМ отклонилось хотя бы на 1. Показатель чувствительности — эмпирическая плотность распределения минимальных расстояний колебаний $\mathbf{X}$, отклоняющих решения ВФМ.

А л г о р и т м. Задается число экспериментов T. Для каждого $D \in \Delta$:

генерируется случайным образом вектор $\mathbf{X}_t'$ (T раз);

для каждого $\mathbf{X}_t'$: вычисляется $D_t' = D(\mathbf{X}_t')$; методом перебора определяется и фиксируется $h_{\mathbf{X}_t'} = \arg\min_{h \in N_0}\{|D_t' - D(\zeta^h(\mathbf{X}_t'))| = 1\}$; на основе статистики по h формируется вектор эмпирических частот $\mathbf{V}_D = (v_1, v_2, \ldots)$, где v_i — отношение числа исходов, когда $h = i$, к числу всех исходов.

Критерий чувствительности. Модель с интерпретацией D обладает хорошей чувствительностью, если для $\mathbf{V}_D$ справедливо: 1) $v_1 < 15\%$; 2) $v_2 < 25\%$.

Комментарии и открытые вопросы. *Выбор количества экспериментов T.* С одной стороны, при большем T достигается большая точность совпадения эмпирических и теоретических статистик. Так как компоненты вектора операционных параметров берутся из конечного подмножества целых неотрицательных чисел, то T не должно быть больше количества таких векторов (также конечного, однако стремительно растущего вместе с увеличением числа факторов и их множеств значений). С другой стороны, с точки зрения практического применения необходимо учитывать вычислительные мощности. Поэтому в качестве эвристического правила можно предложить количество экспериментов, равное произведению трех величин: количества индексов, максимальной мощности множества среди множеств значений их факторов и максимальной мощности подмножества из множеств значений существенных индексов. Так, для трех индексов, каждый из которых зависит от 5 факторов, а они, в свою очередь, оцениваются одним из пяти значений, количество экспериментов составит $T = 3 \cdot 5 \cdot 5^5 = 46875$. Чтобы не перебирать все значения вектора операционных параметров, необходим дальнейший анализ оптимального числа экспериментов T для обоих алгоритмов.

Выбор генератора случайных чисел. Свойство абсолютной случайности для методики не так критично, как возможность рассмотреть максимально различные варианты исходов для изучения свойств модели. При выборе генератора случайных чисел для алгоритма рекомендуются квазислучайные генераторы с хорошим разбросом в пространстве исходов (LDS-последовательности [6]).

Метод перебора одиночных колебаний. Для перебора одиночных колебаний предлагается осуществлять колебания пофакторно: сначала выбирается фактор, определяемый переменной $x_{ij} = x_{11}$; если возможно в рамках допустимых значений N_{ij}, рассматриваются одиночные колебания, образованные изменением заданного фактора на единицу: $x_{ij} \pm 1$ (шаг 1); затем шаг 1 повторяется последовательно для других факторов: $x_{ij} = x_{12}, x_{13}, \ldots, x_{1q}, x_{21}, x_{22}, \ldots, x_{2q}, \ldots, x_{n1}, x_{n2}, \ldots, x_{nq}$. Чтобы уменьшить количество всевозможных переборов одиночных колебаний, необходимо разработать оптимальные методы с использованием случайных последовательностей.

Метод перебора для нахождения $h_{\mathbf{X}_t'}$. Предлагается перебирать колебания в возрастающем порядке количества одиночных колебаний: пусть $h = 1$, $\mathbf{X} = \mathbf{X}_t'$; сначала выбирается фактор, определяемый переменной $x_{ij} = x_{11}$; если возможно в рамках допустимых значений N_{ij}, рассматриваются одиночные колебания, образованные изменением заданного фактора на единицу: $x_{ij} \pm 1$; для таких двух колебаний проверяется условие $|D'(\mathbf{X}) - D(\zeta(\bullet))| = 1$; если оно выполнено, то $h_{\mathbf{X}'_t} = h$ и алгоритм перебора завершает работу; если оно не выполнено, вычисленные отклонения помещаются в очередь ожидания $\mathbf{X}_h$; процесс повторяется последовательно для других факторов: $x_{ij} = x_{12}, x_{13}, \ldots, x_{1q}, x_{21}, x_{22}, \ldots, x_{2q}, \ldots, x_{n1}, x_{n2}, \ldots, x_{nq}$; если алгоритм перебора не остановлен, то полагается $h = h + 1$, и весь процесс повторяется для вынимаемых из очереди ожидания $\mathbf{X}_h$ отклонений. Однако при таком подходе некоторые отклонения проверяются по нескольку раз, поэтому необходимо разработать более оптимальные методы перебора.

Задание параметров для критерия устойчивости и чувствительности. Общим правилом для критерия устойчивости должна быть «малость» колебаний, отклоняющих решение ВФМ более, чем на единицу. Для хорошей чувствительности ВФМ должна менять свое решение при некотором оптимальном числе колебаний, при котором она будет по-разному интерпретировать различные оцениваемые объекты. Однако при этом ВФМ все еще должна быть толерантна к «малым» колебаниям.

Методика в общем случае. Открыт вопрос разработки методики для общего случая, когда проверяются общие свойства модели ВФМ относительно всех входных параметров (операционных и стратегических). Этот тип анализа полезно использовать на стадии внедрения процесса принятия решения, когда осуществляется определение факторов, происходит их агрегирование в однородные группы и специфицируются интерпретации индексов. На основе результатов такого анализа могут быть внесены дополнительные корректировки в конструкцию модели (например, увеличено количество факторов для какого-то индекса). Использование общей методики позволит также понять, насколько модель неустойчива к небольшим отклонениям стратегических параметров. Подразумевается, что изменение стратегических, в отличие от операционных, параметров, конфигурирующих модель, должно вносить более фундаментальные изменения.

Анализ чувствительности и устойчивости в разрезе отдельных факторов. Приведенная методика оперирует не отдельными факторами, а вектором операционных параметров. В случае адаптации алгоритмов для анализа чувствительности и устойчивости ВФМ в разрезе отдельных факторов могут возникать ситуации, когда никакое отклонение фактора в допустимом диапазоне не может привести к изменению решения.

Заключение. Эффективный процесс анализа устойчивости ВФМ требует решения двух задач: формализации ВФМ и автоматизации алгоритмов устойчивости. На основе приведенных в работе алгоритмов разработано программное обеспечение для целей анализа адекватности ВФМ. Применение описанных алгоритмов к модели принятия решений в области ИБ показало его эффективность и практическую значимость [5]. Требуется дальнейшее изучение практических методик анализа ВФМ с целью построения наиболее оптимальной модели с позиции устойчивости для целей ИБ.

СПИСОК ЛИТЕРАТУРЫ

1. *Varian H. R.* How to build an economic model in your spare time. — In: Passion and Craft: Economists at Work./ Ed. by M. Szenberg. Ann Arbor: Univ. Michigan Press, 1997.
2. Managing IT Investments. — Intel Information Technology. 2003, Aug.
3. *Ezingeard J. N.*, *Bowen-Schrire M.* Information Security: A Strategic Issue. Rep. Hanley Management College (UK), Dataforeningen (Sweden), 2003.
4. *Heerkens G. R.* PMP Project Management. N.-Y.: McGraw-Hill, 2005, 266 p.

5. *Кудинов С. В.* Анализ моделей принятия решений по проектам в области информационной безопасности. — Вестник РГГУ, 2009, № 10/09, 178 с.
6. *Sobol I. M.* Uniformly distributed sequences with an additional uniformity property. — USSR Comput. Maths. Math. Phys., 1976, v. 16, p. 236–242.