

УКД 004:685.5

ПРИМЕНЕНИЕ МЕТОДА СТОХАСТИЧЕСКОГО ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ ПРИ ОРГАНИЗАЦИИ КОНТРОЛЯ ПРИМЕНЕНИЯ ОРГАНИЗАЦИОННО-ТЕХНИЧЕСКИХ СИСТЕМ

Волков В.Ф., Пономарев А.С.

*ФГБВОУ ВО «Военно-космическая академия имени А.Ф. Можайского»,
Санкт-Петербург, e-mail: vka@mail.ru*

Рассматриваются два подхода к решению задачи обоснования рационального сочетания операций целевого применения организационно-технической системы (ОТС) и процедур контроля выполнения технико-технологических операций. Первый подход основан на определении параметров корректирующего вектора после каждого шага операции по выполнению целевой задачи. Вероятностной моделью процесса является сложная марковская цепь, вычислительные алгоритмы основаны на выборе варианта по критерию минимума стоимости дополнительно привлекаемых ресурсов при жестком ограничении на продолжительность операции и на вероятность достижения требуемого эффекта. При втором подходе в качестве вероятностной модели рассматривается марковская цепь со случайным моментом остановки; разработан алгоритм выбора управления для каждого исхода каждого этапа процесса применения ОТС. Оптимизационные расчеты как в первом, так и во втором подходе базируются на методе стохастического динамического программирования. Разработанные алгоритмы протестированы на примерах и дают возможность проанализировать влияние различных факторов, входящих в состав комплекса исходных данных (технические характеристики, параметры обстановки, атрибуты способов организации процесса и эксплуатационные характеристики), на степень приближения показателя результативности к максимально возможному для рассматриваемых условий значению.

Ключевые слова: результативность, марковская цепь, пошаговая реконфигурация, ретроспективная развертка, контроль факта исполнения

APPLICATION OF STOCHASTIC DYNAMIC PROGRAMMING METHOD WHILE ORGANIZING THE CONTROL OF ORGANIZATIONAL AND TECHNICAL SYSTEMS

Volkov V.F., Ponomaryov A.S.

Military Space Academy named after A.F. Mozhaysky, St. Petersburg, e-mail: vka@mail.ru

It is supposed two approaches to decision of justification task of the rational combination of organizational and technical systems target applications operations and performing technical and technological operations. First approach is based on parameters of the correction vector determining after each of target application operation performing. Probabilistic model of process is a complex Markov chain, computational algorithms are based on selection a minimum cost of additionally sources option with a strict limit on the duration of the operation and probability of achieving the desired effect. With the second approach is seemed Markov chain with a random moment of stopping as a probabilistic model; the algorithm of control selection for each stage of organizational and technical systems application is developed. Optimization calculations both in the first and in the second approaches are based on stochastic dynamic programming method. Developed algorithms are tested on examples and give the opportunity to analyze the influence of various factors that are part of the source data (technical specifications, environment parameters, attributes of the process organization methods and operational characteristics) on the degree of approximation of the performance indicator to the maximum possible for the conditions under consideration.

Keywords: productivity, Markov chain, step-by-step reconfiguration, retrospective scan, control of the execution fact

Одной из важных частных задач, возникающих при планировании применения сложных организационно-технических систем (ОТС), является установление рационального сочетания непосредственных технико-технологических операций и процедур контроля их выполнения. Очевидно, что дополнительное уточнение факта выполнения целевой задачи, с одной стороны, повышает степень объективности принимаемого решения, но, с другой стороны, требует использования новых ресурсов и сопровождается временными потерями. Возникает новая научная задача по обоснованию рациональной координации целевых воздействий и процедур контроля.

Анализ предметной области в сфере проблем управления в реальных иерархических системах показал, что методы координации

непрерывно развивались в различных работах по исследованию операций. В частности, в работах [1–3] были исследованы два способа организации информационного обмена в процессе координации. Первый предполагает осуществление обмена информацией между центром и подсистемами на каждом шаге итеративного процесса. При втором способе после получения и анализа информации от подсистем центр осуществляет весь итеративный процесс самостоятельно и выработывает управляющий сигнал для подсистем. Первый способ позволяет упростить координирующую задачу, но может зависеть от ограниченных вычислительных мощностей и пропускных способностей каналов связи. Второй способ связан с учетом факторов неопределенности, т.е. центр выраба-

тывает координирующий сигнал, учитывая неопределенность обстановки на основе прогнозирования ее развития; подсистемы действуют и на основе и координирующего сигнала, и реально складывающейся обстановки. Однако в работах [1–3] не предусмотрены вычисления на основе организации ретро-спективной развертки поискового процесса, учитывающие результаты контроля состояния ОТС, ошибки распознавания состояния ОТС, а также не учитываются параметры доверительных интервалов исходной информации (об обстановке). Таким образом, необходимо разработать алгоритм управления, учитывающий результаты контроля исполнения управляющих воздействий. При первом подходе к решению данной задачи наиболее адекватной для расчетов моделью представляется схема вероятностной корректировки плана действий по результатам каждого этапа. При втором подходе в качестве модели целесообразно рассматривать марковскую цепь со случайным моментом останова.

Материалы и методы исследования

Первоначально рассмотрим задачу оптимизации мониторинга результатов применения ОТС на основе пошаговой корректировки плана действий. В методиках и моделях, ориентированных на первый подход, после каждого шага операции по выполнению целевой задачи рассчитываются компоненты корректирующего вектора. Логика оптимизации заключается либо в решении задачи по критерию максимума вероятности достижения требуемого эффекта при жестком ограничении на продолжительность операции и на стоимость дополнительно привлекаемых ресурсов, либо в решении задачи по критерию минимума стоимости дополнительно привлекаемых ресурсов при жестком ограничении на продолжительность операции и на вероятность достижения требуемого эффекта [4; 5]. Однако данные алгоритмы, хотя и отвечают требованию ситуационного реагирования на результаты каждого этапа, не учитывают возросшие возможности перспективных ЭВМ по формированию массива ситуационных решений, соответствующих обстановке. Характеристики компьютеров петафлопсного диапазона и квантовых компьютеров мощностью в несколько кубитов, их оснащение процессорами с архитектурой ARMv9, внедрение алгоритмов Deepfake позволяют осуществлять накапливание и «табулирование» информации по различным исходам каждого шага операции. В связи с этим рассмотрим задачу адаптивного реагирования на результаты этапов при корректировке второго типа, учитывающей реконфигурацию математи-

ческого ожидания и среднеквадратического отклонения продолжительностей t_i не одного, а нескольких последующих этапов (при допущении, что рассматриваемые числовые характеристики величин t_i равны их плановым значениям. Вероятностной моделью процесса совместного прохождения процедур контроля и управляющих воздействий при корректировке второго типа является сложная марковская цепь [6]. При этом необходимо предусмотреть, что если полевые исследования или результаты вычислительных экспериментов подтверждают гипотезу о β -распределении продолжительностей этапов, то для получения значений t_p , используемых в расчетах, необходимо использовать теорему «Об обратной функции» [7]. Особенности «машинных» аспектов реализации соответствующей имитационной модели будут обусловлены «табличной» спецификой β -распределения, и возникающие при этом «разбросы» получаемых оценок требуют минимизации значений ошибок 1-го и 2-го рода (риска пользователя результатов расчетов и риска исполнителей).

При корректировке первого типа последовательно на каждом этапе, включая предпоследний, рассчитывается вектор реконфигурации, изменяющий параметры распределения продолжительности очередного этапа и минимизирующий стоимость дополнительных инвестиций (выбор из нескольких вариантов, предложенных экспертами и проработанных профильными специалистами); при этом отслеживается выполнение требования по гарантированному обеспечению минимума риска срыва выполнения основной задачи. При корректировке второго типа по результатам каждого этапа для всех остальных шагов, включая предпоследний, рассчитывается вектор реконфигурации, параметры которого последовательно уточняются после окончания каждого этапа путем решения соответствующих оптимизационных задач.

Рассмотрим пример пошаговой оптимизации плана действий при корректировке 2-го типа для случая $L = 4$. При фактическом выполнении первого этапа за время t_1^* (например, может иметь место отставание) выражение для продолжительности всей операции уточняется по формуле $T_{\Sigma 1}^* = \sum_{l=2}^4 t_l - t_1^*$. Вероятность выполнения задачи в заданный срок T_{nn} с учетом t_1^* под- лежит пересчету:

$$\gamma_1^* = \frac{1}{2} \left[W \left(\frac{T_{nn} - t_1^* - \bar{t}_2 - \bar{t}_3 - \bar{t}_4}{\sqrt{2} * \sqrt{\sigma_2^2 + \sigma_3^2 + \sigma_4^2}} \right) + 1 \right],$$

где W – табличная функция.

При невыполнении требований руководящих документов (РД), т.е. при $\gamma_1^* < \gamma^{PD}$, необходимо реализовать дополнительные мероприятия, в результате которых произойдет изменение числовых характеристик и второго, и третьего, и четвертого этапов: $m'_2 = m_2 + y_1$; $\sigma'_2 = \sigma_2 + y'_1$; $m'_3 = m_3 + y'_2$; $\sigma'_3 = \sigma_3 + y'_2$; $m'_4 = m_4 + y'_3$; $\sigma'_4 = \sigma_4 + y'_3$. Таким образом, после первого этапа вектор реконфигурации $Y_{<0>} = \langle y_1, y'_1, y_2, y'_2, y_3, y'_3 \rangle$ имеет размерность $\Theta_2 = 6$ и должен удовлетворять ограничению $\Psi_1(y_1, y'_1, y_2, y'_2, y_3, y'_3, t_1, \dots) \geq \gamma^{PD}$ или

$$\frac{1}{2} \left[W \left(\frac{T_{nl} - t_1^* - m_2 - y_1 - m_3 - y_2 - m_4 - y_3}{\sqrt{(\sigma_2 + y'_1)^2 + (\sigma_3 + y'_2)^2 + (\sigma_4 + y'_3) * \sqrt{2}}} \right) + 1 \right] \geq \gamma^{PD}$$

Математическое ожидание затрат на введение в действие обеспечивающих (корректирующих) мероприятий: $\Delta S^{(2)} = \Delta S^{(2)}(x_1, f^{(2)}(y_1, \dots))$, где $x_1 = t_1^* - m_1$, f – зависимость затрат от с параметров $y_1, y'_1, y_2, y'_2, y_3, y'_3$. Пусть второй этап реализован за время $t_2 = t_2^{**} = f(m_2 + y_1; \sigma_2 + y'_1)$. При сохранении отставания и отсутствии временных резервов повторно возникает риск, т.е. при

$$\gamma_2^{**} < \gamma^{PD}, \text{ где } \gamma_2^{**} = \frac{1}{2} \left[W \left(\frac{T_{nl} - t_1^* - t_2^{**} - m_3 - y_2 - m_4 - y_3}{\sqrt{(\sigma_3 + y'_2)^2 + (\sigma_1 + y'_3)^2 * \sqrt{2}}} \right) + 1 \right]$$

осуществляется новая корректировка, вектор реконфигурации имеет размерность $\Theta_2 = 8$:

$$m''_3 = m_3 + y_2 + y_4, \sigma''_3 = \sigma_3 + y'_2 + y'_4; m''_4 = m_4 + y_3 + y_5, \sigma''_4 = \sigma_4 + y'_3 + y'_5$$

Ограничение на вектор реконфигурации: $\Psi_2(y_2, y'_2, y_3, y'_3, y_4, y'_4, y_5, y'_5, t_2^{**}, \dots) = \gamma^{PD}$.

Средние дополнительные затраты от введения в действия мероприятий с параметрами $y_4, y'_4; y_5, y'_5$: $\Delta S^{(3)} = \Delta S^{(3)}(x_2, y_4, y'_4, y_5, y'_5)$, где $x_2 = x_1 + y_1$.

Пусть $t_3 = t_3^{***} = t_3^*(m_3 + y_2 + y_4, \sigma_3 + y'_2 + y'_4)$. При $\gamma_3^{***} < \gamma^{PD}$, где

$$\gamma_3^{***} = \frac{1}{2} \left[W \left(\frac{T_{nl} - t_1^* - t_2^{**} - t_3^{***} - m_4 - y_3 - y_5}{(\sigma_4 + y'_3 + y'_5) * \sqrt{2}} \right) + 1 \right],$$

осуществляется новая корректировка, вектор реконфигурации имеет размерность $\Theta_4 = 6$, при этом

$$M_{i_4} = m_4 + y_3 + y_5 + y_6; \sigma_{i_4} = \sigma_4 + y'_3 + y'_5 + y_1; \text{ т.е. } Y_{<0_4>} = \langle y_3, y'_3, y_5, y'_5, y_6, y'_6 \rangle^T = Y_{<6>}$$

Ограничение на вектор реконфигурации: $\gamma_3^{***} \geq \gamma^{PD}$, где

$$\gamma_3^{***} = \frac{1}{2} \left[W \left(\frac{T_{nl} - t_1^* - t_2^{**} - t_3^{***} - m_4 - y_3 - y_5 - y_6}{(\sigma_4 + y'_3 + y'_5 + y'_6) * \sqrt{2}} \right) + 1 \right],$$

или $\Psi_3(y_3, y'_3, y_4, y'_4, y_5, y'_5, y_6, y'_6, t_1^*, t_2^{**}, t_3^{***}, \dots) \geq \gamma^{PD}$. Средняя стоимость вводимых резервных ресурсов y_6, y'_6 : $\Delta S^{(4)} = \Delta S^{(4)}(x_3, y_5, y'_6)$, где $x_3 = x_2 + y_2 + y_4$.

Суммарная стоимость всех дополнительно задействуемых ресурсов:

$$\begin{aligned} \eta_{2,3,4}(x_1, y_1, y'_1, y_2, y'_2, y_3, y'_3, y_4, y'_4, y_5, y'_5, y_6, y'_6) = \\ = \Delta S^{(2)}(x_1, f^{(2)}(y_1, \dots)) + \Delta S^{(3)}(x_2, y_4, y'_4, y_5, y'_5) + \Delta S^{(4)}(x_3, y_5, y'_6). \end{aligned}$$

Таким образом, требуется определить такие значения параметров $\tilde{y}_1, \tilde{y}'_1, \tilde{y}_2, \tilde{y}'_2, \tilde{y}_3, \tilde{y}'_3, \tilde{y}_4, \tilde{y}'_4, \tilde{y}_5, \tilde{y}'_5, \tilde{y}_6, \tilde{y}'_6$, при которых показатель $\eta_{2,3,4}$ обращается в минимум и выполняются ограничения:

$$\begin{aligned} \Psi_1(t_1^*, y_1, y'_1, y_2, y'_2, y_3, y'_3) &\geq \gamma^{PD}; \\ \Psi_2(t_1^*, t_2^{**}, y_2, y'_2, y_3, y'_3, y_4, y'_4, y_5, y'_5) &\geq \gamma^{PD}; \\ \Psi_3(y_3, y'_3, y_4, y'_4, y_5, y'_5, y_6, y'_6, t_1^*, t_2^{**}, t_3^{***}, \dots) &\geq \gamma^{PD}. \end{aligned}$$

Сформулированная задача решается инерционно с оперативными издержками (временем на сбор информации для регрессионного уточнения коэффициентов зависимости расходов на реконфигурацию от «подбираемых» параметров корректировки).

Результаты исследования и их обсуждение

Для проведения анализа результатов исследования рассмотрим процесс оптимизации мониторинга на основе координации управляющих воздействий и процедур контроля на примере схемы, в которой решение о завершении целенаправленного применения ОТС принимается по результатам проведения контроля факта исполнения управляющей программы [8; 9]. Введем обозначения:

c_l – стоимость ресурса, расходуемого на одну попытку целевого воздействия;

h – вероятность реализации одношаговой операции целевого воздействия (например, по переводу из одного штатного режима в другой);

a_l – стоимость операции контроля состояния ОТС;

B – показатель выходного эффекта (выигрыш от успешного штатного функционирования ОТС; предотвращенный ущерб и т.п.).

Возможны следующие случайные исходы операции: Z_1 – задача выполнена с первой попытки; Z_2 – процесс проходил по схеме «неудачная попытка → повторная попытка → контроль → удачная попытка → контроль»; Z_3 – процесс проходил по схеме «неудачная попытка → контроль → неудачная попытка → контроль → удачная попытка → контроль» и т.д. Исходное множество комбинаций целевых воздействий и процедур контроля определяется экспертами с учетом вероятностно-психологической модели принятия решений (модели действий) операторов.

Для перехода из каждого «нецелевого» состояния (результатов наступления событий Z_2, Z_3, Z_4, \dots) требуется плановый (заранее рассчитанный) расход того или иного ресурса. Так как вероятность перехода процесса из состояния в состояние зависит от результатов предыдущего исхода, то вероятностной моделью рассматриваемого процесса является марковская цепь со случайным моментом останова [6].

Методика решения задачи включает 3 этапа. Первый этап: генерирование исходного множества альтернатив путем проведения предварительной экспертизы структурных, параметрических и организационных атрибутов вариантов на предмет реализуемости. Второй этап: формирование множе-

ства вариантов методом ретроспективной развертки и заполнения базы аналитических решений соответствующих рекуррентных уравнений. Третий этап: определение конкретных комбинаций чередования целевых воздействий и процедур контроля результатов воздействий.

Из вышеизложенного следует, что выигрыш на любом l -м этапе зависит от типа управления. Введем обозначения: $K^{(1)}$ – первый тип управления (воздействие проводится без последующей процедуры распознавания состояния ОТС); $K^{(2)}$ – второй тип (после каждого сеанса проводится контроль факта перехода ОТС в требуемый режим); $K^{(3)}$ – третий тип (проведение управляющих воздействий и процедур контроля прекращается по решению руководства). Найдём связь между средним выигрышем V при нахождении системы в состоянии $U^{(j)}$, $j = 0(1)(L-1)$ и типом управления. Для первого типа управления $V_{jj+1} = -C_l + B(1-h)^j h + 0[1 - (1-h)^j h]$, но возможен подход, когда вместо нуля подставляются штрафные вычеты. Для второго типа $V_{jj+1}(K^{(2)}) = -c_l - a_l + B(1-h)^{j+1} + 0_*[1 - (1-h)^{j+1}]$. Выигрыш для третьего типа управления приравняем к нулю: $V_{jj+1}(K^{(3)}) = 0$. Далее, применив принцип проактивного управления для многоэтапной оптимизации и учитывая требование независимости от локальных экстремумов, сформулируем правила согласованного управления для каждого этапа и каждого состояния. В соответствии с логикой учёта возможности любого хода событий первоначально просчитаем выигрыш на последнем этапе. Система после $(L-1)$ -го этапа может находиться либо в состоянии $U^{(L)}$ (цель достигнута), либо в состояниях $U^{(L-1)}, U^{(L-2)}, \dots, U^{(1)}$, для которых возможна реализация двух типов управления – $K^{(1)}$ и $K^{(3)}$. При применении управления $K^{(1)}$ средний выигрыш на последнем этапе равен

$$V_l(U^{(j)}, K^{(1)}) = B(1-h)^j h - a_l, j = 0(1)(L-1).$$

После алгебраических преобразований получаем рекуррентные выражения для ретроспективной развертки. Если результат $(n-1)$ -го этапа – состояние $U^{(L)}$, то необходимым управляющим воздействием является $K^{(3)}$, т.е. выигрыш на всех оставшихся этапах равен нулю. Если результат $(n-1)$ -го этапа – состояние $U^{(j)}$, $j = 0(1)(n-1)$, то при управлении $K^{(1)}$ выигрыш рассчитывается по формуле

$$\begin{aligned} \bar{V}_n(U^{(j)}, K^{(1)}) + \bar{V}_{n+1, \dots, L}(U^{(j)}, K^{(1)}) = \\ = B(1-h)^j * h - c_l + \bar{V}_{n+1, \dots, L}(U^{(j+1)}). \end{aligned}$$

При управлении $K^{(2)}$ в этой же ситуации функциональное уравнение имеет вид

$$\bar{V}_{n,\dots,L}(U^{(j)}, K^{(2)}) = B(1-h)^j * h - c_l - \\ - a_l + (1-h)^{(j+1)} \bar{V}_{(n+1),\dots,L}^*(U^{(0)}).$$

При управлении $K^{(3)}$, если после $(n-1)$ -го этапа наступило состояние $U^{(j)}$, $j = 0(1)(n-1)$, средний выигрыш на всех остальных этапах будет равен нулю. Если по результатам вычислений получаются одинаковые выигрыши, то для отыскания единственного решения необходимо рассчитать среднеквадратическое отклонение как корень второй степени из математического ожидания квадрата центрированного значения выигрыша.

Рассмотрим реализацию алгоритма на примере дистанционного управления и контроля перевода в требуемый режим гипотетической ОТС при следующих исходных данных (в относительных единицах): $L = 6$; $h = 0.3$; $c_l = 1$; $a_l = 0.4$; $B = 8$. После вычислений получаем рациональную комбинацию для каждого исхода каждого этапа:

$$K_6^* : K^{(1)} \rightarrow K^{(1)} \rightarrow K^{(1)} \rightarrow K^{(3)} \rightarrow K^{(3)} \rightarrow K^{(3)},$$

$$K_5^* : K^{(1)} \rightarrow K^{(2)} \rightarrow K^{(2)} \rightarrow K^{(3)} \rightarrow K^{(3)},$$

$$K_4^* : K^{(2)} \rightarrow K^{(2)} \rightarrow K^{(2)} \rightarrow K^{(3)},$$

$$K_3^* : K^{(2)} \rightarrow K^{(2)} \rightarrow K^{(2)},$$

$$K_2^* : K^{(1)} \rightarrow K^{(2)},$$

$$K_1^* : K^{(1)},$$

при этом максимальный выигрыш составит $V^* = 3.03$. Следует отметить, что при направленном переборе массива параметров исходной информации предлагаемый алгоритм дает возможность проанализировать влияние различных факторов, входящих в состав комплекса исходных данных (соотношение величин L ; h ; c_p ; a_p ; B), на степень приближения показателя результативности V^* к исходному значению B .

Заключение

Разработаны алгоритмы решения задачи по обоснованию рациональной координации целевых воздействий в процессе применения организационно-технической системы и процедур контроля их результатов. При первом подходе рассматривается сложная марковская цепь со степенью зависимости исходов этапов, определяемой возможностями своевременного создания массива уточненной информации для выбранной схемы. При втором подходе в качестве вероятностной модели рассматривается марковская цепь со случайным моментом остановки; корректировка плана действий по результатам каждого этапа основана на методе стохастического динамического программирования.

Разработанные алгоритмы протестированы на примерах пошаговой корректировки четырехэтапного и шестиэтапного планов (аналитическое и численное решения).

Список литературы

1. Микони С.В., Соколов Б.В., Юсупов Р.М. Квалиметрия моделей и полимодельных комплексов: монография. М.: РАН, 2018. 314 с.
2. Новиков Д.А. Аналитическая сложность решения задач управления организационно-техническими системами // Автоматика и телемеханика. 2018. № 5. С. 107–118.
3. Моисеев В.С. Групповое применение беспилотных летательных аппаратов: монография. Казань: Редакционно-издательский центр «Школа», 2017. 572 с.
4. Баранов В.В., Цвиркун А.Д. Управление развитием: структурный анализ, задачи, устойчивость // Автоматика и телемеханика. 2018. № 10. С. 55–75.
5. Пшихопов В.Х., Медведев М.Ю. Групповое управление движением мобильных роботов в неопределенной среде с использованием неустойчивых режимов // Труды СПИИРАН. 2018. № 5(60). С. 39–60.
6. Андреев В.Н., Волков В.Ф., Иоффе А.Я. Анализ оптимальных решений в марковских процессах с доходами // Вопросы кибернетики, вып. 74. М.: АН СССР, 1981. С. 33–42.
7. Мадера А.Г. Метод определения вероятностей прогнозируемых событий при принятии решений // Искусственный интеллект и принятие решений. 2016. № 2. С. 38–45.
8. Волков В.Ю. Адаптивная многопороговая селекция объектов на изображениях в системах дистанционного зондирования // Информационно-управляющие системы. 2020. № 3. С. 12–25.
9. Григорьев А.Н., Попович Т.В. Принципы и применение использования технологий дистанционного зондирования в информационном обеспечении // Труды ВКА имени А.Ф. Можайского. Выпуск 664. СПб.: ВКА имени А.Ф. Можайского, 2018. С. 51–60.