

Моделирование многоуровневой архитектуры системы адаптивного управления реагированием пожарно-спасательных подразделений

Р. Ш. Хабибулин

Академия государственной противопожарной службы МЧС России

kh-r@yandex.ru

Аннотация. Представлена формализованная модель многоуровневой системы адаптивного управления реагированием пожарно-спасательных подразделений. Процесс реагирования формализован как полумарковский процесс принятия решений, вложенный в теоретико-множественное описание системы. Обоснована задача адаптивного выбора критерия оптимальности по оперативному контексту пожара. Предложена пятиуровневая архитектура управления с расширенным адаптивным циклом. Описана платформа симуляционного моделирования на основе обучения с подкреплением, верифицированная на сценарии тушения пожара резервуарного парка.

Ключевые слова: адаптивное управление; пожарно-спасательные подразделения; полумарковские процессы; мета-оптимизация; обучение с подкреплением; цифровой двойник

I. ВВЕДЕНИЕ

Пожарно-спасательные подразделения (ПСП) функционируют в условиях высокой динамичности обстановки на пожаре, неполноты информации и критической ответственности за результат. Традиционные модели управления, основанные на жёстких регламентах и централизованном принятии решений, демонстрируют недостаточную эффективность при нестандартных условиях [1, 2]. Существующие системы поддержки принятия решений (СППР) в основном решают отдельные задачи, однако не обеспечивают сквозной формализованной иерархии управления [3, 4].

Целью работы является разработка формального описания многоуровневой архитектуры системы адаптивного управления реагированием ПСП, обоснование процедуры адаптивного выбора критерия оптимальности и экспериментальная верификация архитектуры на платформе симуляционного моделирования.

II. ФОРМАЛЬНАЯ МОДЕЛЬ

Определение 1. Система адаптивного управления реагированием определяется как кортеж:

$$(L, A, R, S, C, F, G, T) \quad (1)$$

где S – пространство состояний обстановки на пожаре (ситуационное пространство); R – множество ресурсов; A – множество агентов системы управления; $L = \{l_1, l_2, \dots, l_5\}$ – уровни иерархии; C – каналы информационного обмена; F – функция адаптации; G – целевое

пространство (множество допустимых исходов); T – временная ось.

Определение 2. Ситуационное пространство: $S = \{s \mid s = \langle O, E, H, W, \tau \rangle\}$, где O – характеристики объекта защиты; E – параметры пожара; H – угрозы жизни людей; W – условия среды; $\tau \in T$ – временная метка.

Определение 3. Функция адаптации отображает текущее состояние и отклонение в корректирующее воздействие:

$$\Phi: S \times R \times G \rightarrow \Delta\pi \quad (2)$$

Условие активации адаптации на уровне l_k : $\delta s_k = |s(\tau) - s^*(\tau)| > \varepsilon_k$, где ε_k – порог чувствительности уровня. На каждом уровне определён коэффициент автономности $\alpha_k \in [0, 1]$. Функция адаптации F корректирует α_k в зависимости от отклонения δs_k и состояния каналов связи: при $\delta s_k > \varepsilon_k$ и недоступности восходящего канала s_k выполняется $\alpha_k \rightarrow \min(\alpha_k + \Delta\alpha, 1, 0)$, обеспечивая переход к нижестоящему уровню управлению.

Определение 4. Коэффициент автономности $\alpha_k \in [0, 1]$ характеризует степень самостоятельности уровня l_k при принятии решений: $\alpha_k = 0$ означает полную зависимость от вышестоящего уровня, $\alpha_k = 1$ – полную автономию. Значение α_k динамически корректируется функцией адаптации F в зависимости от состояния каналов связи и величины отклонения δs_k .

Определение 5. Политика управления $\pi: S \times X \rightarrow A$ – отображение, ставящее каждой паре (состояние обстановки, оперативный контекст) допустимое тактическое решение. Оптимальная политика π^* минимизирует выбранный критерий $f \in \{f_1, \dots, f_5\}$: $\pi^* = \arg \min f(\pi)$.

Функционирование отдельного уровня $l_k \in L$ формализуется как полумарковский процесс принятия решений (ПМПР) [14, 15]:

$$M_k = (S_M, A_{M(s)}, P, D\tau, X, C_M) \quad (3)$$

где $S_M = \{s_0, s_1, \dots, s_5\}$ – конечное множество фаз реагирования (от обнаружения пожара до ликвидации последствий), конкретизирующее ситуационное пространство S кортежа (1) для оперативного уровня; $A_{M(s)}$ – множество допустимых тактических решений в фазе s ; $D\tau(\cdot|s, a, x)$ – распределение времени в фазе реагирования (гамма-распределение по эмпирическим данным [14]); X – пространство оперативного контекста пожара $x = (x_{stat}, x_{dyn}, x_{info})$.

В большинстве существующих моделей критерий оптимальности задаётся фиксированным. Анализ пожаров объектов топливно-энергетического комплекса (ТЭК) показал, что коэффициент вариации времени фаз локализации и ликвидации горения составляет 3,3–4,2, что в 4–5 раз превышает пороговое значение, при максимальном наблюдаемом времени локализации 813 мин [14]. Это делает минимизацию среднего времени реагирования неадекватной для затяжных пожаров.

Для решения этой проблемы сформулирована задача мета-оптимизации: найти отображение $\kappa^*: X \rightarrow \Omega_{ctrl}$, сопоставляющее каждому оперативному контексту на пожаре критерий, обеспечивающий наилучшие результаты [16]. Пространство критериев разделено на два класса: критерии управления Ω_{ctrl} (минимум среднего времени реагирования f_1 , взвешенное суммарное время реагирования f_2 , нормативный квантиль f_3 , риск-чувствительный f_4 [16, 17]) и информационный критерий f_5 (минимум запаздывания данных мониторинга пожара).

Выбор критерия реализуется системой приоритетных правил: (κ_3) при запаздывании мониторинга выше допустимого – f_5 ; (κ_2) при коэффициенте вариации времени реагирования $\geq 0,75$ – риск-чувствительный f_4 ; (κ_1) в базовом режиме – f_1 или f_2 .

III. МНОГОУРОВНЕВАЯ АРХИТЕКТУРА И АДАПТИВНЫЙ ЦИКЛ

Архитектура адаптивного управления включает пять иерархических уровней (рисунок 1), каждый из которых характеризуется функциональным составом F_k , коэффициентом автономности α_k и метрикой эффективности μ_k . Задача мета-оптимизации κ^* (раздел 2) образует мета-уровень, логически предшествующий работе всех пяти уровней: оптимальная организационная структура управления на пожаре и политика (стратегия) управления не могут быть синтезированы без предварительного определения критерия оптимизации.

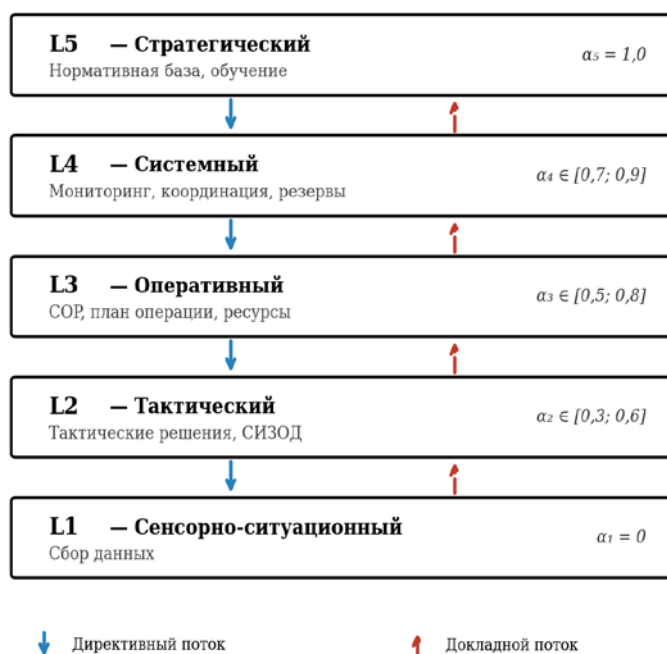


Рис. 1. Иерархическая архитектура системы адаптивного управления

Уровень L1 – сенсорно-ситуационный. Сбор первичных данных (АПС, БПЛА и др.) и формирование ситуационной картины (СОР). $\alpha_1 = 0$.

Уровень L2 – тактический. Уровень командира отделения. Tактические решения, контроль, перераспределение личного состава. $\alpha_2 \in [0,3; 0,6]$.

Уровень L3 – оперативный (штаб пожаротушения). Ведение единой оперативной картины, планирование операции, управление ресурсами, прогноз обстановки. $\alpha_3 \in [0,5; 0,8]$.

Уровень L4 – системный (ЦУКС). Мониторинг гарнизона, межгарнизонная координация, резервирование ресурсов. $\alpha_4 \in [0,7; 0,9]$.

Уровень L5 – стратегический. Нормативная база, институциональное обучение, долгосрочное планирование. $\alpha_5 = 1,0$.

Для каждой пары смежных уровней (l_k, l_{k+1}) определяются нисходящий директивный $c_{\downarrow k}$ и восходящий докладной $c_{\uparrow k}$ каналы. На уровнях L2–L3 реализуется горизонтальная координация между параллельно действующими отделениями.

На каждом уровне иерархии реализуется расширенный адаптивный цикл НОРД – «Наблюдение – Ориентация – Решение – Действие» (аналог OODA-loop [7]), дополненный пятым этапом «Контроль» для обеспечения обратной связи (рис. 2). Период исполнения цикла соответствует горизонту управления уровня.

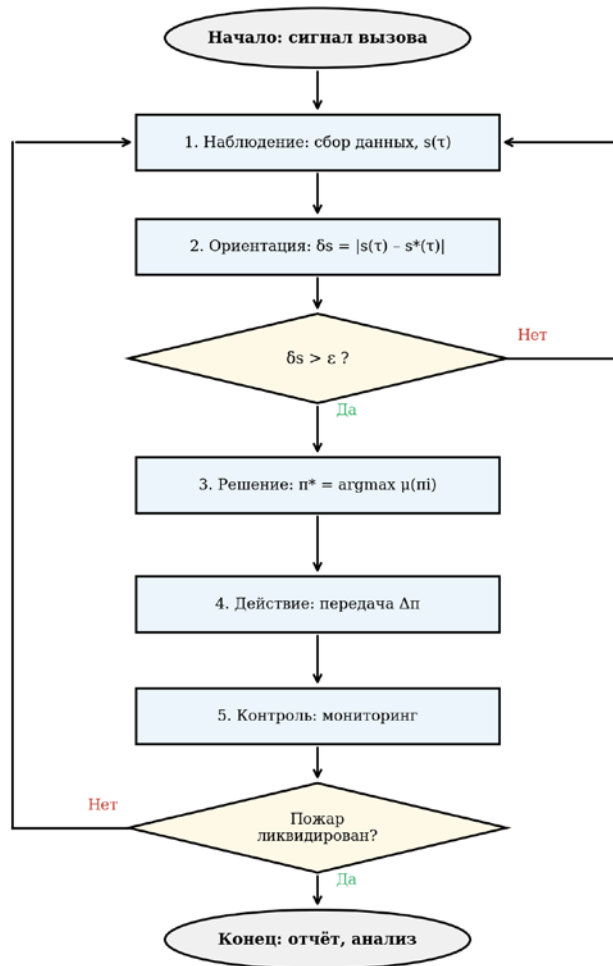


Рис. 2. Алгоритм расширенного адаптивного цикла НОРД

Этапы цикла: (1) Наблюдение – сбор данных, обновление $s(t)$; (2) Ориентация – вычисление отклонения δ_s и проверка условия адаптации; (3) Решение – генерация альтернатив и выбор оптимального π^* по критерию, определённым правилами κ_1 – κ_3 ; (4) Действие – передача управляющих воздействий $\Delta \pi$ нижестоящим уровням; (5) Контроль – мониторинг исполнения, при повторном отклонении – возврат к наблюдению. По итогам цикла НОРД на уровне $L3$ оперативный контекст x обновляется, что может привести к пересмотру κ^* и смене критерия оптимизации (например, переход от f_1 к f_4 при возрастании коэффициента вариации).

ТАБЛИЦА I. Соответствие трёх иерархий в модели адаптивного управления

Организационная (L1–L5)	Мета-оптимизация ($\kappa \rightarrow O \rightarrow \pi$)	RL-агент платформы
L5 Стратегический	выбор κ^*	RL-L3 (32×3)
L4 Системный	синтез O^*	RL-L2 (64×5)
L3 Оперативный	оптимизация π^*	RL-L1 (256×15)
L2 Tактический	исполнители π^*	Среда (DES)
L1 Сенсорный	данные $\rightarrow X$	Среда (DES)

IV. ПЛАТФОРМА СИМУЛЯЦИОННОГО МОДЕЛИРОВАНИЯ

Из рассмотренных подходов к моделированию (агентное, дискретно-событийное, гибридное, цифровой двойник [9, 10]) для реализации выбран гибридный подход – агентное управление с обучением с подкреплением, сочетающий потоковую симуляцию физических процессов с автономностью агентов. На его

основе разработана платформа «SAUR PSP v. 1.0» (рис. 3).

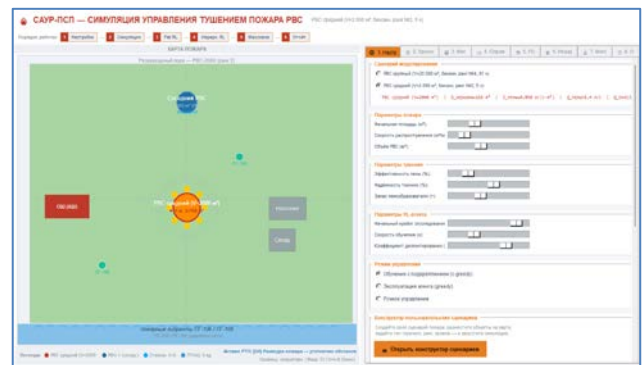


Рис. 3. Интерфейс компьютерной программы SAUR PSP v. 1.0

Программный комплекс (Система Адаптивного Управления Реагированием Пожарно-Спасательных Подразделений) представляет собой интерактивную среду имитационного моделирования процессов тушения пожаров на объектах хранения нефтепродуктов с интегрированными алгоритмами обучения с подкреплением.

Система предназначена для:

- моделирования динамики развития и тушения пожаров в резервуарных парках;

- исследования эффективности адаптивных стратегий управления силами и средствами;
- обучения и оценки компетенций руководителей тушения пожаров (РТП);
- сравнительного анализа обычных и иерархических алгоритмов обучения с подкреплением в задачах оперативного управления.

Приложение поддерживает три режима:

Тренажёр. Пользователь (обучаемый РТП) пошагово выбирает действия; система оценивает каждое решение по Q -таблице обученного агента, начисляя баллы (от +20 за оптимальное до -5 за недопустимое). По завершении сценария генерируется отчёт с пошаговым разбором решений.

СППР. Агент автоматически рекомендует действия; пользователь может принять рекомендацию или переопределить её. Отчёт фиксирует степень согласия человека и агента.

Исследовательский режим. Полный доступ к обучению обычного и иерархического агентов, экспорту данных для отчета.

В каждом режиме моделируется динамика площади горения $S_{fire}(t)$, расход огнетушащих веществ $Q(t)$, число стволов тушения и охлаждения $N(t)$ и интегральный индекс риска $R(t)$.

Управление тушением осуществляется агентом на базе метода Q -learning [11]. Реализованы два режима: обычный одноуровневый Q -learning (единая Q -таблица 256×15) и иерархический [13], воспроизводящий вертикаль адаптивного управления (таблица). Иерархический агент включает три уровня: $L3$ (стратегический, 32×3), $L2$ (тактический, 64×5) и $L1$ (оперативный, 256×15). Составная функция награды нижнего уровня включает интринзическую компоненту, поощряющую согласованность действий с макроцелями вышестоящего уровня управления.

V. РЕЗУЛЬТАТЫ ВЕРИФИКАЦИИ

Верификация проведена на сценарии тушения пожара резервуара вертикального стального (РВС) объёмом 20 000 м³ с бензином (ранг пожара №4). Коэффициент вариации времени локализации для данного класса объектов составляет 3,3, что значительно превышает порог 0,75, поэтому по правилу κ_2 выбран риск-чувствительный критерий $f_4 = CVaR_\alpha(T)$. Функция награды агента включает штраф за индекс риска $R(t)$, что соответствует минимизации критерия f_4 .

Результаты сравнительного эксперимента: иерархический агент снизил максимальный индекс риска на 13,8% (с 1,00 до 0,86), расход огнетушащих веществ – на 5,4% (с 650 до 615 л/с), время ликвидации – с 640 до 510 мин по сравнению с обычным агентом.



Рис. 4. Результаты имитационного моделирования

Полученные результаты подтверждают, что иерархическая декомпозиция управления, соответствующая реальной вертикали управления, превосходит одноуровневый подход по ключевым метрикам. Перспективным направлением является замена табличного Q -learning на $Deep Q$ -Network для непрерывного пространства состояний, увеличение обучающей базы с ретроспективными данными о пожарах и интеграция с геоинформационными системами.

VI. ЗАКЛЮЧЕНИЕ

В работе предложена формализованная модель многоуровневой системы адаптивного управления реагированием пожарно-спасательных подразделений в виде кортежа $\langle L, A, R, S, C, F, G, T \rangle$. Функционирование каждого уровня формализовано как полумарковский процесс принятия решений, вложенный в кортеж. Обоснована задача мета-оптимизации – адаптивного выбора критерия оптимальности (f_1 – f_5) по оперативному контексту пожара. Предложена пятиуровневая архитектура управления с расширенным адаптивным циклом НОРД и динамическим коэффициентом автономности. Разработана платформа имитационного моделирования «SAUR PSP v. 1.0», верификация которой на сценарии тушения пожара РВС показала, что иерархический агент превосходит обычный одноуровневый по метрикам риска (-13,8%) и времени ликвидации (-20,3%), что подтверждает практическую значимость предложенной архитектуры управления.

СПИСОК ЛИТЕРАТУРЫ

- [1] Брушлинский Н.Н. Моделирование оперативной деятельности пожарной службы. М.: Стройиздат, 1981. 95 с.
- [2] Денисов А.Н., Порошин А.А., Данилов М.М. и др. Генезис развития, современные реалии научно-методического и нормативного правового обоснования понятий «крупный пожар», «сложный (затяжной) пожар» // Пожаровзрывобезопасность. 2025. Т. 34, №2. С. 5–19.
- [3] Брушлинский Н.Н. Системный анализ деятельности ГПС. М.: МИПБ МВД России, 1998. 255 с.
- [4] Абдурагимов Г.И., Таранцев А.А. Теория массового обслуживания в управлении пожарной охраной. М.: Академия ГПС МВД России, 2000. 101 с.

- [5] Матюшин А.В., Порошин А.А., Матюшин Ю.А. Отечественный опыт расчётного обоснования ресурсов оперативных подразделений пожарной охраны // Пожарная безопасность. 2005. №3. С. 61–74.
- [6] Степанов Е.В., Чан М.Х.Х., Гринченко Б.Б. и др. Модель и методика оценки степени сложности системы управления ПСП на пожаре // Современные проблемы гражданской защиты. 2023. №1(46). С. 47–56.
- [7] Brehmer B. The Dynamic OODA Loop: Amalgamating Boyd's OODA Loop and the Cybernetic Approach to Command and Control // Proc. 10th International Command and Control Research and Technology Symposium. McLean, VA, 2005.
- [8] Wooldridge M. An Introduction to MultiAgent Systems. 2nd ed. Chichester: Wiley, 2009. 484 p.
- [9] Macal C.M., North M.J. Tutorial on agent-based modelling and simulation // Journal of Simulation. 2010. Vol. 4, No. 3. P. 151–162.
- [10] Grieves M., Vickers J. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems // Transdisciplinary Perspectives on Complex Systems. Springer, 2017. P. 85–113.
- [11] Watkins C.J.C.H., Dayan P. Q-learning // Machine Learning. 1992. Vol. 8, No. 3–4. P. 279–292.
- [12] Sutton R.S., Barto A.G. Reinforcement Learning: An Introduction. 2nd ed. Cambridge: MIT Press, 2018. 552 p.
- [13] Dietterich T.G. Hierarchical Reinforcement Learning with the MAXQ Value Function Decomposition // Journal of Artificial Intelligence Research. 2000. Vol. 13. P. 227–303.
- [14] Хабибулин Р.Ш. Прогнозирование временных интервалов реагирования пожарно-спасательных подразделений на пожары объектов топливно-энергетического комплекса на основе гамма-распределений и марковской цепи // Моделирование сложных процессов и систем: Сборник трудов XXXV Международной научно-практической конференции, Химки, 26 февраля 2025 года. – Химки: Академия гражданской защиты Министерства Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий имени генерал-лейтенанта Д.И. Михайлика, 2025. С. 82–86. – EDN VZQVJS.
- [15] Puterman M.L. Markov Decision Processes: Discrete Stochastic Dynamic Programming. New York: Wiley, 1994. 678 p.
- [16] Rockafellar R.T., Uryasev S. Conditional value-at-risk for general loss distributions // Journal of Banking and Finance. 2002. Vol. 26. P. 1443–1471.
- [17] Chow Y., Ghavamzadeh M., Janson L., Pavone M. Risk-constrained reinforcement learning with percentile risk criteria // Journal of Machine Learning Research. 2017. Vol. 18, No. 167. P. 1–51.