

## МОДЕЛЬ УПРАВЛЕНИЯ ЗАБОРОМ ТОПЛИВА В ТРАНСПОРТНОМ ХОЛДИНГЕ

**Цыганов В.В.**

*Институт проблем управления им. В.А. Трапезникова РАН, Москва, Россия*  
bbc@ipu.ru

*Аннотация. В модели управления транспортным холдингом, забор топлива контролируется менеджером холдинга (Регулятором). Холдинг имеет сервисное бюро, которое возглавляет Руководитель. Ему помогает Эксперт. Этому бюро подчинено подразделение, которое непосредственно обслуживает клиентов. Это подразделение возглавляет Интендант. Ни Руководитель, ни Эксперт не знают минимально необходимый забор топлива в этом подразделении. Пользуясь этим, Интендант может манипулировать забором топлива, чтобы получить большее поощрение. Подобным образом, Руководитель может манипулировать забором топлива в бюро, чтобы получить более высокую категорию. Найдены достаточные условия, при которых и Интендант, и Руководитель стремятся минимизировать забор топлива.*

*Ключевые слова: Транспорт, холдинг, забор топлива, управление, обучение, манипулирование.*

### Введение

При организации и управлении обслуживанием железнодорожного транспорта широко используется концепция цифрового предприятия [1]. В её основе – использование искусственный интеллект (ИИ) и машинного обучения (МО) [2]. МО позволяет изучать и управлять сложными производственными системами, без необходимости более глубокого понимания и математического моделирования [3]. Спектр приложений, использующих МО, постоянно расширяется [4]. Однако применение МО в моделях управления затруднено из-за отсутствия соответствующей теоретической базы [5].

С другой стороны, при использовании МО в управлении крупным холдингом (таким, как ОАО «РЖД»), важно учитывать человеческий фактор. Система управления таким холдингом – многоуровневая. При этом руководители низших звеньев, более близких к производству и оказанию услуг, лучше знают их возможности, чем руководители высшего уровня. В таких случаях в теории управления организационными системами [6] говорят об асимметричной осведомленности элементов системы. Проблемы взаимодействия информированных и неинформированных элементов организационной системы анализировались в [7]. При этом более информированные элементы системы имеют возможность использовать недостаточную информированность остальных для достижения собственных целей.

Поэтому теоретические исследования и практическая отработка методов и процедур МО в крупном холдинге должны учитывать активность его менеджеров – возможность манипулирования информацией, чтобы влиять на результаты МО в свою пользу [8]. Теоретическое исследование такого рода проблем, в условиях динамики и неопределенности, проводится с позиций теории активных систем (АС) [9,10]. При этом используется агентное моделирование (agent-based modeling), которое применяется, в частности, на железнодорожном транспорте [11].

Быстро развивающееся направление теории АС — синтез механизмов функционирования АС с элементами ИИ и МО (кратко – интеллектуальных механизмов). Основные классы интеллектуальных механизмов – механизмы функционирования АС с процедурами адаптации (кратко – адаптивные механизмы, см., например [12]) и с процедурами цифрового (машинного) обучения (кратко – обучающиеся механизмы [13]).

Важный подкласс обучающихся механизмов составляют механизмы функционирования АС, в которых используются процедуры цифрового самообучения (кратко – самообучающиеся механизмы [14]). Другой подкласс обучающихся механизмов составляют механизмы функционирования АС, в котором используются процедуры цифрового обучения с учителем (учебные механизмы [15]).

Подкласс более сложных гибридных обучающихся механизмов функционирования АС, использующих различные сочетания процедур цифрового обучения, включает двухканальные, консультативные и экспертные механизмы. Давно известный вид гибридных обучающихся механизмов - двухканальные механизмы [16]. Другой их вид – консультативные механизмы, в которых учителя используют процедуры цифрового самообучения [17]. В отличие от них, в экспертных механизмах учителя не используют процедуры цифрового самообучения [18]. Содержательно, экспертный механизм – это механизм с процедурами цифрового обучения, в котором учитель является знающим (сложившимся) экспертом, у которого нет текущей необходимости самообучаться [19].

В данной статье рассматривается гибридный обучающийся механизм, использующий комбинации цифровых процедур, используемых в обучающихся и экспертных механизмах.

## 1. Структура модели управления забором топлива в холдинге

Рассмотрим модель управления забором топлива в холдинге с трехуровневой структурой. В целом, забор топлива регулирует ответственный менеджер холдинга (кратко - Регулятор). Для оказания профильных услуг, в холдинг имеется сервисное бюро, возглавляемое Руководителем. Ему подчинено подразделение, непосредственно обслуживающее клиентов холдинга. Это подразделение возглавляет Интендант.

На забор топлива в бюро влияют случайные внешние воздействия. Поэтому у Регулятора нет полной информации о требуемом для бюро заборе топлива. Чтобы уменьшить забор топлива, Регулятор использует информацию, полученную от Руководителя. Однако Руководитель, пользуясь неосведомленностью Регулятора, может манипулировать забором топлива на нужды бюро и подчиненного ему подразделения, чтобы повлиять на решения Регулятора и увеличить собственное вознаграждение. Соответственно, Регулятору необходимо научиться управлять Руководителем так, чтобы забор топлива в бюро был минимальным при любых случайных внешних воздействиях.

Со своей стороны, Руководитель, в момент принятия решения, знает минимально необходимый забор топлива для нужд бюро. Но он не знает минимально возможного забора топлива подчиненного ему подразделения из-за случайных внешних воздействий.

Наконец, Интендант, в момент принятия решения, знает минимально необходимый подразделению забор топлива. Таким образом, он может манипулировать этим забором топлива, чтобы влиять на решения Руководителя относительно его собственного вознаграждения. Соответственно, Руководителю необходимо одновременно учиться и управлять Интендантом таким образом, чтобы уменьшить забор топлива. Чтобы делать это, Руководитель использует помощь Эксперта.

## 2. Управление Интендантом

### 2.1. Обучение Руководителя

Введем следующие обозначения:

- $t$  – дискретный период времени,  $t=0,1,\dots$ ;
- $F_t$  – забор топлива подразделением в периоде  $t$ .

Величину  $F_t$  Интендант сообщает Руководителю и Эксперту. Нижний предел забора топлива подразделением равен  $f_t: F_t \in P_t = [f_t, \beta]$ . Здесь  $f_t$  — стационарная случайная величина:  $f_t \in P = [\alpha, \varphi]$ ,  $\alpha > 0$ ,  $\varphi \leq \beta$ .

Руководитель и Эксперт не знают  $f_t$ . Однако значение  $f_t$  становится известно Интенданту в периоде  $t$ . Таким образом, Интендант может манипулировать  $F_t$ , если  $F_t \in P_t = [f_t, \beta]$ . Соответственно, Руководитель стремится уменьшить забор топлива Интендантом в условиях асимметричной информированности [6]. Для этого Руководитель оперирует относительными показателями (индикаторами) забора топлива:

$$i_t = f_t/\beta, i_t \in [\alpha/\beta, 1], I_t = F_t/\beta, I_t \in I_t = [i_t, 1]. \quad (1)$$

Таким образом, Руководитель должен добиться, чтобы Интендант снизил показатель забора топлива  $I_t$  до нижнего предела  $i_t$ . С этой целью, Руководитель присваивает Интенданту номинацию 1 за успешное снижение забора топлива, и поощряет его. И, наоборот, за перерасход топлива, Интендант получает номинацию 0, при которой поощрение отсутствует.

Чтобы правильно номинировать Интенданта в условиях неопределенности, Руководитель должен учиться. Предположим, что для этого Руководитель использует процедуру цифрового обучения с учителем, в роли которого выступает Эксперт. Именно, в периоде  $t$  Руководитель использует заключение Эксперта  $e_t(I_t)$  о целесообразности показателя забора топлива  $I_t$  Интендантом [9]. Если Эксперт считает забор топлива в периоде  $t$  избыточным, то  $e_t(I_t) = 1$ , в противном случае –  $e_t(I_t) = 0$ . Формально,

$$e(I_t) = \begin{cases} 1, & \text{если } I_t > b \\ 0, & \text{если } I_t \leq b \end{cases}, t=0,1,\dots, \quad (2)$$

где  $b$  – параметр заключения. Эксперт сообщает свое заключение  $e_t(I_t)$  Руководителю для номинирования.

Неправильная номинация приводит к потерям. Если  $I_t$  является стационарной случайной величиной, то обучение Руководителя принятию решений о номинации Интенданта проводится так, чтобы минимизировать средние потери. Следуя [9], обозначим:

- $h_{10}$  – потери при несправедливой номинации 0 (хотя Интендант заслуживает номинации 1);
- $h_{01}$  – потери, возникающие из-за не заслуженной Интендантом номинации 1.

Настраиваемая оценка параметра номинации в периоде  $t$ , позволяющая минимизировать средние потери (кратко – стандарт  $s_t$ ), вычисляется с помощью алгоритма обучения с учителем [9]:

$$s_{t+1} = S(s_t, I_t) = s_t - \lambda_t [s_t - 0.5 - h_{01} + (h_{01} + h_{10})e_t(I_t)], \quad s = s^0, \quad (3)$$

где  $S(s_t, I_t)$  – процедура стандартизации,  $0 < \lambda_{t+1} < \lambda_t$ ,  $\sum_{\tau=0}^{\infty} \lambda_{\tau} < \infty$ ,  $t=0,1,\dots$ . Используя (3), Руководитель определяет номинацию Интенданта ( $d_t$ ) с помощью соответствующей процедуры:

$$d_t = D(s_t, I_t) = \begin{cases} 1, & \text{если } I_t \leq s_t \\ 0, & \text{если } I_t > s_t \end{cases} \quad (4)$$

Процедура стандартизации (3) и процедура номинации (4) образуют систему обучения Руководителя проведению номинации и управлению Интендантом с помощью Эксперта (кратко – экспертную систему)  $M = (D, S)$ .

## 2.2. Решения Интенданта

Предположим, что полезность Интенданта растет при увеличении текущей и будущих номинаций:

$$\Sigma_t = \Sigma[d_t, d_{t+1}, \dots, d_{t+\phi}], \quad \Sigma_{\tau} \uparrow d_{\tau}, \tau = \overline{t, t + \phi}, \quad t=0,1,\dots \quad (5)$$

где  $\Sigma[\bullet]$  – неубывающая функция,  $\phi$  – дальновидность Интенданта. Чтобы увеличить полезность (5), Интендант выбирает в каждом периоде  $t$  забор топлива  $F_t$ , характеризуемый его индикатором  $I_t(1)$ . Предположим, что:

- Интендант знает только то, что  $f_t \in P$  и  $F_t \in P_{\tau}$ ,  $\tau = \overline{t + 1, t + \phi}$ ;
- при устранении этой неопределенности Интендант руководствуется принципом максимального гарантированного результата [7].

Тогда целью Интенданта является максимизация целевой функции  $V_t(M, I_t)$ , определяемой как минимальное гарантированное значение полезности (5):

$$V_t(M, I_t) = \min_{\tau=\overline{t+1, t+\phi}} \min_{f_{\tau} \in P} \min_{F_{\tau} \in P_{\tau}} \Sigma_{\tau} \xrightarrow{I_t \in I_t} \max. \quad (6)$$

Оптимальное для Интенданта решение  $F_t$  максимизирует целевую функцию (6). Таким образом, совокупность оптимальных решений Интенданта:

$$K_t(M) = \{I_t^* | V_t(M, I_t^*) \geq V_t(M, I_t), I_t \in I_t\}. \quad (7)$$

Будем предполагать гипотезу благосклонности Интенданта к Руководителю: если  $i_t \in K_t(M)$ , то  $I_t = i_t$ ,  $t=0,1,\dots$

## 2.3. Экспертная система Руководителя

Цель Руководителя - уменьшение забора топлива  $F_t$  до нижнего предела  $f_t$ . Согласно (4), номинация Интенданта растет, когда показатель  $I_t$  снижается до стандарта  $s_t$ . Это мотивирует Интенданта на снижение показателя  $I_t$ .

С другой стороны, на практике, будущий стандарт забора топлива обычно снижается, когда фактический забор топлива снижается до текущего стандарта. Формально, в нашей модели, это означает, что стандарт  $s_{t+1}$  в следующем периоде ( $t+1$ ) снижается при снижении показателя забора топлива  $I_t$  до текущего стандарта  $s_t$ .

Согласно (4), чем ниже стандарт  $s_t$ , тем меньший показатель забора топлива  $I_{t+1}$  потребуется в периоде  $t+1$  для повышения номинации Интенданта. Это не мотивирует дальновидного Интенданта к снижению показателя забора топлива  $I_t$  ниже стандарта  $s_t$ . Это проблема типична для планирования от достигнутого уровня [9].

**Утверждение.** Экспертная система  $M = (D, S)$  достаточна для мотивации Интенданта к снижению забора топлива в каждом периоде, так что  $F_t = f_t$ ,  $t=0,1,\dots$

**Доказательство.** Экспертная система  $M = (D, S)$  включает процедуры (2)-(4). Согласно (5), целевая функция Интенданта (6) растет как с текущей, так и с будущими номинациями  $d_{\tau}$ ,  $\tau = \overline{t, t + \phi}$ . Согласно (4), текущая номинация  $d_t = D(n_t, I_t)$  не уменьшается с уменьшением  $I_t$ .

Рассмотрим зависимость будущей номинации  $d_\tau = D(s_\tau, I_\tau)$  от  $I_t$ ,  $\tau = \overline{t+1, t+\phi}$ . Согласно (2),  $e_\tau(I_\tau)$  в периоде  $\tau$  не растет с уменьшением  $I_t$ . Тогда, согласно (3), стандарт  $s_\tau$  не снижается с уменьшением  $I_t$ . Кроме того, согласно (4), номинация  $d_\tau = D(s_\tau, I_\tau)$  не уменьшается при уменьшении  $s_\tau$ . Следовательно, номинация  $d_\tau = D(s_\tau, I_\tau)$  в периоде  $\tau$  не уменьшается при уменьшении  $I_t$ ,  $\tau = \overline{t+1, t+\phi}$ . Таким образом, из (5) и (6) получаем, что целевая функция  $V_t(M, I_t)$  не уменьшается при уменьшении  $I_t$ . Так как  $I_t \geq i_t$  согласно (4), то  $V_t(M, i_t) \geq V_t(M, I_t)$ . Таким образом, из (7)  $i_t \in K_t(M)$ , и, согласно гипотезе о доброжелательности Интенданта к Руководителю,  $I_t^* = i_t$ . Тогда из (4) следует, что  $F_t = f_t$ ,  $t = 0, 1, \dots$ , ч.т.д.

Содержательно, Утверждение означает, что для того, чтобы заинтересовать Интенданта в снижении забора топлива, Руководителю достаточно использовать экспертную систему  $M = (D, S)$ .

### 3. Управление забором топлива в бюро

#### 3.1. Взаимодействие Регулятора и Руководителя

В рамках гипотезы об асимметричной осведомленности активных элементов, будем предполагать, что Регулятор не знает забора топлива  $F_t$  (о котором Интендант уже успел сообщить Руководителю в периоде  $t$ ). Кроме того, Регулятор не знает нижнего предела забора топлива для нужд бюро ( $w_t$ ), где  $w_t$  – стационарная случайная величина:  $w_t \in \Lambda = [\varepsilon, \delta]$ ,  $\varepsilon > 0, \delta \leq \varepsilon$ .

Чтобы повысить свою осведомленность, Регулятор требует от Руководителя информации о заборе топлива Интендантом ( $E_t$ ), а также о заборе топлива для нужд бюро ( $W_t$ ) в периоде  $t$ . Здесь  $E_t \in \Phi_t = [F_t, \gamma]$ ,  $\gamma \geq \beta$ ,  $W_t \in \Lambda_t = [w_t, \delta]$ .

С другой стороны, в начале периода  $t$ , Руководителю становятся известны значения  $F_t$  и  $w_t$ . Это позволяет ему манипулировать информацией о заборах топлива в своих интересах, сообщая Регулятору  $E_t$  и  $W_t$  вместо  $F_t$  и  $w_t$ . Соответственно, Регулятору необходимо уменьшить  $E_t$  и  $W_t$ .

Напомним, что нижний предел забора топлива, о котором Интендант может сообщить Руководителю, составляет  $f_t$ , так что  $F_t \geq f_t$ ,  $t=0, 1, \dots$ . Но Регулятор не знает нижнего предела забора топлива  $f_t$ . Несмотря на это, Регулятору необходимо обеспечить  $E_t = f_t$ ,  $W_t = w_t$ .

Рассмотрим, какие задачи для этого должен решить Регулятор. Забор топлива  $E_t$ , сообщаемый ему Руководителем, согласно (8), зависит от информации, которую Руководитель получает от Интенданта (т.е. от  $F_t$ ). Следовательно, Регулятор должен, во-первых, убедиться, что Руководитель не преувеличивает сообщаемый забор топлива  $E_t$  (т.е.  $E_t = F_t$ ).

Во-вторых, Регулятор должен добиться того, чтобы Руководитель поощрял Интенданта минимизировать забор топлива, сообщаемый Руководителю (т.е.  $F_t = f_t$ ). Для этого Регулятор делегирует Руководителю право устанавливать систему управления Интендантом  $Z$ ,  $Z \in \Theta$ . Здесь  $\Theta$  – множество систем управления Интендантом, которые Регулятор разрешает использовать Руководителю. Кроме того, Регулятор должен следить за тем, чтобы Руководитель не завышал забор топлива для собственных нужд  $W_t$  (т.е.  $W_t = w_t$ ).

#### 3.2. Полный забор топлива Бюро

Сообщаемый Регулятору полный забор топлива бюро  $B_t$  в периоде  $t$  равен сумме забора топлива  $E_t$ , о котором Руководитель сообщает Регулятору, и забору топлива для собственных нужд  $W_t$ :  $B_t = E_t + W_t$ . Основываясь на потреблении топлива  $F_t$ , ранее сообщенном Интендантом, Руководитель может сообщить Регулятору о нижнем пределе полного забора топлива в бюро  $b_t$ :  $b_t = F_t + w_t$ . Таким образом,  $b_t \in B_t = [F_t + w_t, \gamma + \delta]$ ,  $B_t \subset B = [\alpha + \varepsilon, \gamma + \delta]$ .

Регулятор стремится сократить полный забор топлива бюро  $B_t$ , о котором сообщил Руководитель, причем  $B_t \geq b_t$ . Из  $b_t = F_t + w_t$  и  $F_t \geq f_t$  следует  $B_t \geq f_t + w_t$ . Обозначим  $y_t \equiv f_t + w_t$ . Поскольку  $f_t$  и  $w_t$  являются стационарными случайными величинами, то и  $y_t$  также является стационарной случайной величиной,  $y_t \in Y = [\alpha + \varepsilon, \varphi + \delta]$ . Таким образом,  $B_t \in B_t = [b_t, \beta + \delta]$ .

По сути,  $y_t$  – это минимальный полный забор топлива в бюро. Согласно гипотезе асимметричной осведомленности [6], мы предполагаем, что Регулятор не знает  $y_t$ . Но значение  $b_t$  становится известно Руководителю в периоде  $t$  до того, как будет выбрано  $B_t$ . Таким образом, Руководитель может манипулировать забором топлива в свою пользу, выбирая  $B_t$ ,  $B_t \geq y_t$ . Такие манипуляции часто происходят при иерархическом управлении холдингом [9,12].

### 3.3. Самообучение Регулятора

Регулятор должен мотивировать Руководителя на достижение минимума забора топлива бюро  $y_t$ :  $B_t = y_t$ ,  $t = 0, 1, \dots$ . Для этого Регулятор должен пройти процедуру обучения и управления, назначая Руководителю вознаграждение за минимизацию забора топлива бюро  $B_t$ . Предположим, что это вознаграждение связано с категорией, которую Регулятор присуждает Руководителю на основании его результатов. А именно, Регулятор присваивает Руководителю первую категорию, если забор топлива разумный, и вторую категорию, если он избыточный. Руководитель поощряется только в случае присвоения первой категории.

Таким образом, действуя в условиях неопределенности, Регулятор должен отнести Руководителя к одну из двух категорий (первой или второй). Сначала рассмотрим случай, когда Регулятору известен нижний предел полного забора топлива в бюро ( $y_t$ ). Чтобы определить категорию Руководителя в периоде  $t$ , Регулятору необходимо присвоить  $y_t$  одному из двух подмножеств  $Y_1$  и  $Y_2$ , составляющих множество  $Y$ ,  $Y_1 \cup Y_2 = Y$ . Если  $y_t$  относится к подмножеству  $Y_1$ , то Руководитель получает категорию 1, в противном случае – категорию 2.

Неправильная категоризация приводит к убыткам. Задача состоит в том, чтобы определить подмножества  $Y_1$  и  $Y_2$ , при которых средние потери категоризации минимальны. Поскольку  $y_t$  является стационарной случайной переменной, мы можем использовать подход, развитый в [12]. А именно, для каждого неизвестного множества  $Y_1$  и  $Y_2$ , введем функции потерь:

- $L_1(u, y_t) = y_t - tu$  – потери в случае, если Регулятор считает, что  $y_t \in Y_1$ , и Руководитель получает первую категорию (т.е. Регулятор считает забор топлива разумным, в то время как на самом деле забор топлива избыточен:  $y_t \in Y_2$ ,  $0 < m < 1$ );
- $L_2(u, y_t) = l(u - y_t)$  – потери в случае, если Регулятор считает, что  $y_t \in Y_2$ , и Руководитель получает категорию 2 (т.е. Регулятор считает забор топлива Руководителем избыточным, а на самом деле  $y_t \in Y_1$ , т.е. забор топлива - разумный),  $l > 0$ .

Здесь  $u$  – неизвестный параметр, настраиваемая оценка которого  $u_{t+1}$  в периоде  $t+1$  определяется из рекуррентного уравнения [12]:

$$u_{t+1} = U(u_t, y_t) = \begin{cases} u_t + \xi_t m, & \text{если } y_t \leq h_t \\ u_t - \xi_t l, & \text{если } y_t > h_t \end{cases}, \quad u_0 = u^0, \quad h_t = \frac{u_t(m+l)}{l+1}, \quad \sum_{\tau=1}^{\infty} \xi_{\tau} < \infty. \quad (8)$$

Предположим теперь, что Регулятор не знает  $y_t$ . Тогда, для оценки  $k_{t+1}$  величины  $u_{t+1}$ , он может использовать формулу (8), в которой неизвестная величина  $y_t$  заменяется известной величиной  $B_t$ :

$$k_{t+1} = U(k_t, B_t) = \begin{cases} k_t + \xi_t m & \text{если } B_t \leq v_t \\ k_t - \xi_t l, & \text{если } B_t > v_t \end{cases}, \quad k_0 = u^0, \quad v_t = \frac{k_t(m+l)}{l+1}. \quad (9)$$

Процедура (9) называется процедурой обучения Регулятора. Зная  $v_t$ , Регулятор определяет категорию Руководителя:

$$c_t = C(v_t, B_t) = \begin{cases} 1, & \text{если } B_t \leq v_t \\ 2, & \text{если } B_t > v_t \end{cases}. \quad (10)$$

Процедура  $C(v_t, B_t)$  называется процедурой категоризации. По сути,  $v_t$  – это норма забора топлива, используемая Регулятором для определения категории Руководителя: если забор топлива  $B_t$  не больше этой нормы ( $B_t \leq v_t$ ), то Руководитель получает первую категорию ( $c_t = 1$ ), в противном случае – вторую категорию ( $c_t = 2$ ). Процедура обучения (9) и процедура категоризации (10) объединены в систему управления забором топлива Руководителем с категоризацией (сокращенно, систему категоризации)  $\Omega = (C, E)$ .

### 3.4. Решения Руководителя

Учитывая систему категоризации  $\Omega = (C, E)$ , Руководитель принимает решения, которые обеспечивают ему более высокую полезность. Решения Руководителя касаются:

- системы управления Интендантом  $Z$ ,  $Z \in \Theta$ , где  $\Theta$  – множество систем управления Интендантом, которые Регулятор разрешает использовать Руководителю;
- сообщаемого Регулятору забора топлива Интендантом:  $E_t$ ,  $E \in \Phi_t$ ;
- сообщаемого Регулятору забора топлива для нужд бюро  $W_t$ ,  $W_t \in \Lambda_t$ .

Руководитель принимает эти решения ( $Z, E_t, W_t$ ) на основе полезности, зависящей от текущей и будущих категорий в периоде  $t$ :

$$G_t = G[c_t, c_{t+1}, \dots, c_{t+\pi}], \quad G_t \downarrow c_j, \quad j = \overline{t, t+\pi}. \quad (11)$$

где  $G[\bullet]$  — невозрастающая функция категорий сегодня и в будущем,  $\pi$  — дальновидность Руководителя.

Предположим, что Руководитель знает, что  $f_t \in P$ ,  $F_t \in P_t$ ,  $w_t \in \Lambda$ ,  $W_t \in \Lambda_t$ ,  $\tau = \overline{t+1, t+\pi}$ . Для устранения этой неопределенности, Руководитель руководствуется принципом максимально гарантированного результата [7]. При этом целевая функция Руководителя  $\Pi_t(Z, F_t, W_t)$  определяется как нижняя граница гарантированного значения полезности (11):

$$\Pi_t(Z, F_t, W_t) = \min_{\tau=\overline{t+1, t+\pi}} \min_{w_t \in \Lambda} \min_{W_t \in \Lambda_t} \min_{f_t \in P} \min_{F_t \in P_t} G_t \xrightarrow{Z \in \Theta, E_t \in \Phi_t, W_t \in \Lambda_t} \max. \quad (12)$$

Руководитель выбирает  $Z, E_t, W_t$  так, чтобы максимизировать целевую функцию (12). Таким образом, совокупность решений Руководителя

$$\text{Argmax}_{Z \in \Theta, E_t \in \Phi_t, W_t \in \Lambda_t} \Pi_t(Z, E_t, W_t) = \{Z^*, E_t^*, W_t^* | \Pi_t(Z^*, E_t^*, W_t^*) \geq \Pi_t(Z, E_t, W_t), Z \in \Theta, E_t \in \Phi_t, W_t \in \Lambda_t\}. \quad (13)$$

### 3.5. Минимизация забора топлива

**Теорема.** Предположим, что:

- экспертная система  $M = (D, S)$  принадлежит множеству систем управления Интендантом  $\Theta$ , которые Регулятор разрешает применять Руководителю:  $M \in \Theta$ ;
- Руководитель поддерживает Регулятора: если  $(M, f_t, r_t) \in \text{Argmax}_{Z \in \Theta, E_t \in \Phi_t, W_t \in \Lambda_t} \Pi_t(Z, E_t, W_t)$ , то

Руководитель выбирает  $Z^* = M$ ,  $E_t^* = f_t$ ,  $W_t^* = w_t$ ,  $t=0,1,\dots$

- экспертная система  $M = (D, S)$  и система категоризации  $\Omega = (C, E)$  объединены в систему управления забором топлива в холдинге (кратко, холдинговую систему)  $H = (M, \Omega)$ .

Тогда для минимизации полного забора топлива бюро в каждом периоде:  $B_t = x_t$ ,  $t = 0,1, \dots$ , Регулятору достаточно использовать холдинговую систему  $H = (M, \Omega)$ .

**Доказательство.** По условию теоремы, холдинговая система  $H = (M, \Omega)$  включает систему категоризации  $\Omega = (C, E)$  с процедурами (9) и (10). Согласно (10), по мере роста  $B_t$ ,  $c_t = C(v_t, B_t)$  не уменьшается. Далее, из (9) следует, что  $c_\tau$  не растет с уменьшением  $B_t$ ,  $\tau = \overline{t+1, t+\pi}$ . Но, согласно (10),  $c_\tau$  не увеличивается с ростом  $k_\tau$ . Таким образом,  $c_\tau = C(v_\tau, B_\tau)$  не увеличивается с ростом  $B_t$ . Следовательно, согласно (11),  $G_t$  не убывает с ростом  $B_t$  при  $f_t \in P$ ,  $F_t \in P_t$ ,  $w_t \in \Lambda$ ,  $W_t \in \Lambda_t$ ,  $\tau = \overline{t+1, t+\pi}$ . Таким образом, согласно (12),  $\Pi_t(Z, E_t, W_t)$  не увеличивается с ростом  $B_t$ .

Из  $B_t = E_t + W_t$  следует, что  $\Pi_t(Z, E_t, W_t)$  не растет с увеличением как  $E_t$ , так и  $W_t$ , при любых  $f_t \in P$ ,  $F_t \in P_t$ ,  $w_t \in \Lambda$ , и  $W_t \in \Lambda_t$ ,  $\tau = \overline{t+1, t+\pi}$ . Следовательно,  $\Pi_t(Z, E_t, W_t)$  достигает максимума при  $E_t = F_t$ , так что  $W_t = w_t$ . Согласно (13),  $(Z, F_t, w_t) \in \text{Argmax}_{E_t \geq D_t, W_t \in \Lambda_t} \Pi_t(Z, E_t, W_t)$ ,  $B_t = F_t + w_t$ ,  $d\Pi_t(Z, E_t, W_t^*) = \Pi_t(Z, F_t, w_t)$ .

Исследуем  $\Pi_t(Z, F_t, w_t)$ , как функцию  $F_t$ . Согласно (10), по мере роста  $F_t$ ,  $c_t = C(r_t, B_t)$  не уменьшается. Далее, из (9) следует, что  $c_\tau$  не растет с уменьшением  $F_t$ ,  $\tau = \overline{t+1, t+\pi}$ . Но, согласно (10),  $c_\tau$  не растет с увеличением  $k_\tau$ . Таким образом,  $c_\tau = C(v_\tau, B_\tau)$  не растет с увеличением  $F_t$ ,  $\tau = \overline{t+1, t+\pi}$ . Следовательно, согласно (11),  $G_t$  не растет с увеличением  $F_t$  при  $f_t \in P$ ,  $F_t \in P_t$ ,  $w_t \in \Lambda$ , и  $W_t \in \Lambda_t$ ,  $\tau = \overline{t+1, t+\pi}$ . Таким образом, согласно (12),  $\Pi_t(Z, F_t, w_t)$  не растет с увеличением  $F_t$  при этих параметрах. Тогда согласно (13)  $(Z, f_t, w_t) \in \text{Argmax}_{F_t \geq f_t, R_t \in \Lambda_t} \Pi_t(Z, E_t, W_t)$ .

Таким образом, для достижения максимума  $\Pi_t(Z, F_t, w_t)$  по  $F_t$ , Руководителю достаточно обеспечить  $F_t = f_t$ ,  $t = 0,1, \dots$ . В свою очередь, для этого, согласно Утверждению, Руководителю достаточно использовать экспертную систему  $M = (D, S)$ . Таким образом,  $\Pi_t(Z, F_t, w_t)$  достигает максимума по  $F_t$  при  $Z = M$ .

Суммируя сказанное выше при доказательстве теоремы, получаем, что  $\Pi_t(Z, E_t, W_t)$  достигает максимума при  $Z = M$ ,  $F_t = f_t$  и  $W_t = w_t$ . Согласно (13),  $(M, f_t, w_t) \in \text{Argmax}_{Z \in \Theta, E_t \in \Phi_t, R_t \in \Lambda_t} \Pi_t(Z, E_t, W_t)$ . Но, согласно Теореме, холдинговая система  $H = (M, \Omega)$  включает в себя экспертную систему  $M = (D, S)$ , причем  $M = (D, S) \in \Theta$ . Тогда, учитывая поддержку Регулятора Руководителем, последний выбирает  $Z^* = M$ ,  $E_t^* = f_t$ ,  $W_t^* = w_t$ ,  $t=0,1,\dots$ . В этом случае,  $B_t = E_t^* + W_t^* = f_t + w_t = y_t$ ,  $t = 0,1, \dots$ , ч.т.д.

Отметим, что для уменьшения забора топлива в бюро, Регулятору достаточно использовать систему категоризации  $\Omega = (C, E)$ . Данная система заинтересовывает Руководителя:

- построить экспертную систему  $M = (D, S)$ , заинтересовывающую Интенданта в минимальном заборе топлива:  $F_t = f_t$ ;
- не превышать забор топлива в подразделении:  $E_t = F_t = f_t$ ;
- использовать возможность минимизировать забор топлива для собственных нужд:  $W_t = w_t$ .

#### 4. Пример: управление забором топлива для обслуживания вагонов

Рассмотрим пример применения развитого подхода к управлению забором топлива в рамках разработки системы управления обслуживанием вагонов, в соответствии с концепцией Digital Enterprise [1]. В [8] для такой разработки использовалось мультиагентное моделирование. В [13] указанная система рассматривается, как многоканальная.

Следуя подходу, развитому в [1,8,13], рассмотрим структуру управления обслуживанием вагонов в крупном холдинге (типа ОАО «РЖД»), как иерархию с тремя менеджерами (как агентами), связанными четырьмя прямыми и обратными каналами. Управляющий холдингом этой службы является прототипом Регулятора. Прототипом Руководителя можно считать неопытного (начинающего) начальника бюро обслуживания вагонов, обучающегося с помощью Эксперта, и подчиняющегося Регулятору. Руководителю, в свою очередь, подчинено подразделение, непосредственно обслуживающее вагоны, и возглавляемое Интендантом.

В условиях неопределенности, Регулятор должен учиться и поощрять Руководителя за снижение забора топлива в бюро. Как следует из сказанного в п. 3.5, чтобы заинтересовать Руководителя в снижении забора топлива, Регулятору достаточно использовать систему категоризации  $\Omega = (C, E)$ .

Соответственно, Регулятор ежемесячно контролирует полный забор топлива бюро при обслуживании вагонов. Если  $t$  — номер месяца, то  $B_t$  — сообщаемый Регулятору в этом месяце полный забор топлива в бюро. Далее, Регулятор рассчитывает норму забора топлива  $v_t = k_t(m + l)/(l + 1)$ , где  $k_t$  вычисляется по (9). Затем, используя (10), Управляющий определяет категорию Руководителя  $c_t = C(v_t, B_t)$ . В случае, если Руководителю присвоена категория 1, он получает премию. Такая система категоризации побуждает Руководителя:

- учиться с помощью Эксперта;
- использовать экспертную систему  $M = (D, S)$ , заинтересовывающую Интенданта в минимальном заборе топлива:  $F_t = f_t$ ;
- не превышать забор топлива подразделением:  $E_t = F_t$ ;
- минимизировать забор топлива для собственных нужд:  $W_t = w_t$ .

Этот пример иллюстрирует простоту и прозрачность развитого подхода к управлению забором топлива в холдинге с трехуровневой структурой управления.

#### 5. Заключение

Построен гибридный механизм, использующий комбинацию цифровых процедур обучающегося и экспертного механизмов для минимизации забора топлива в холдинге с трехуровневой структурой управления, в условиях неопределенности, и с учетом человеческого фактора. В перспективе предполагается использование развитого подхода к управлению забором топлива в крупномасштабных иерархических организациях с использованием более сложных процедур обучения.

#### Литература

1. *Bannikov D., Sirina N.* Service maintenance and repair of passenger cars in the concept of Digital Enterprise // IOP Conf. Series: Materials Science and Engineering. – 2020. – Vol. 91. – P. 012168.
2. *Савушкин С.А., Горбунов В.Г., Цыганов В.В.* Интеллект в цифровой транспортной системе // Информационные технологии в науке, образовании и управлении. 2018. N. 4. – С. 40–44.
3. *Sutton R.* Reinforcement learning: an introduction / Sutton R., Barto A. – Massachusetts: The MIT Press, 2nd edition, 2018.
4. *Bertsekas D.* Reinforcement learning and optimal control. – Athena: Athena Scientific, 2019.
5. *Recht B.* Reflections on the learning-to-control renaissance // Proc. of the 21st World Congress of IFAC. – Berlin, 2020. – P. 4707.
6. *Burkov V.* Mechanism synthesis and management / Burkov V., Kondratiev V., Korgin N., Novikov D. – N.-Y: NOVA Publishers, 2013.
7. *Auster S.* Asymmetric awareness and moral hazard // Games and Economic Behavior. – 2013. – Vol. 82. – P. 503–521.

8. *Цыганов В.В.* Механизмы функционирования транспортной корпорации // Труды Всеросс. совещания по проблемам управления (ВСПУ-2019) – М.: ИПУ РАН, 2019. – С.1836-1841.
9. *Щепкин А.В., Кондратьев В.В., Цыганов В.В.* Механизмы функционирования активных систем в условиях динамики и неопределенности // Теории активных систем – 50 лет: труды междунар. конф. – М.: ИПУ РАН, 2019. – С. 49–54.
10. *Джафаров Э.И., Цыганов В. В.* Противозатратный механизм функционирования трехуровневой активной системы // 15-я мультikonф. по проблемам управления (МКПУ-2022): материалы конф. Математическая теория управления и ее приложения (МТУИП-2022). – СПб.: Концерн ЦНИИ Электроприбор, 2022. – С.93-96.
11. *Alexandrov A., Bannikov D., Sirina N.* Agent-based modeling of service maintenance and repair of rolling stock // IOP Conference Series: Earth and Environmental Science. – 2019. – Vol. 403. – P. 012193.
12. *Цыганов В.В.* Адаптивные процедуры нормирования энергозатрат на железной дороге // Управление развитием крупномасштабных систем (MLSD'2020): труды Тринадцатой междунар. конф.: в 2 т. – М.: ИПУ РАН, 2020. – Т. 2. – С. 1129-1137.
13. *Басыров С.К., Ковалев С.М., Цыганов В.В.* Механизм обучения энергоэффективности на цифровом железнодорожном транспорте // Интеллектуальные системы управления на железнодорожном транспорте. Компьютерное и математическое моделирование (ИСУЖТ-2019): труды Восьмой научно-технич. конф. — М.: НИИАС, 2019. – С. 209–215.
14. *Роцин А.А., Цыганов В.В.* Самообучение ранжированию предприятия железнодорожного транспорта по энергоэффективности // Вестник Уральского гос. университета путей сообщения. 2020. N 2. – С. 69–80.
15. *Федянин Д.Н., Цыганов В.В.* Механизмы наставничества в организационных системах и повышение энергоэффективности железнодорожного транспорта // Вестник Уральского гос. университета путей сообщения. 2019. N 3. – С. 55–66.
16. *Еналеев А.К., Саматов Р.А.* Анализ и синтез механизмов управления в двухканальных активных системах // Системы управления и информационные технологии. 2016. N 4. – С. 28–34.
17. *Цыганов В. В.* Обучение дихотомическому ранжированию и управлению в двухуровневой активной системе с советником // 15-я мультikonф. по проблемам управления (МКПУ-2022): материалы конф. Математическая теория управления и ее приложения (МТУИП-2022). – СПб.: Концерн ЦНИИ Электроприбор, 2022. С.121-124.
18. *Цыганов В. В.* Механизмы контролируемого обучения в крупномасштабном транспортном холдинге // Управление развитием крупномасштабных систем (MLSD'2021): труды Четырнадцатой междунар. конф: в 2 т. – М.: ИПУ РАН, 2021. – Т. 2. –С.1098-1105.
19. *Цыганов В. В.* Обучение с подкреплением и управление иерархическим вагоноремонтным производством // Управление развитием крупномасштабных систем (MLSD'2022): труды Пятнадцатой междунар. конф: в 2 т. – М.: ИПУРАН, 2022. – Т. 2. – С.834-840.
20. *Bannikov D., Sirina N., Smolyaninov A.* Model of the maintenance and repair system in service maintenance management // Transport problems. – 2018. – Vol. 13, N 3. – P. 5-14.