

**Институт проблем управления  
им. В.А. Трапезникова РАН**

# **УПРАВЛЕНИЕ БОЛЬШИМИ СИСТЕМАМИ**

***Выпуск 90  
Март 2021***

**СБОРНИК  
ТРУДОВ**

ISSN 1819-2467

Регистрационный номер Эл. №ФС77-44158 от 09 марта 2011 г.

**Москва – 2021**

РОССИЙСКАЯ АКАДЕМИЯ НАУК  
*Институт проблем управления*  
*им. В.А. Трапезникова*

**УПРАВЛЕНИЕ  
БОЛЬШИМИ  
СИСТЕМАМИ**

**СБОРНИК ТРУДОВ**

*Выпуск 90*

**Москва – 2021**

### КООРДИНАЦИОННЫЙ СОВЕТ

Академики РАН: Васильев С.Н., Емельянов С.В., Куржанский А.Б., Федосов Е.А., Черноусько Ф.Л.; члены-корреспонденты РАН: Желтов С.Ю., Каляев И.А., Пархоменко П.П., Попков Ю.С.; д-ра техн. наук: Кузнецов О.П., Кульба В.В., Лотоцкий В.А., Павлов Б.В., Поляк Б.Т., Рутковский В.Ю.

### РЕДАКЦИОННАЯ КОЛЛЕГИЯ

**Главный редактор:** член-корр. РАН Новиков Д.А. **Зам. главного редактора:** д-р физ.-мат. наук Губко М.В. **Отв. секретарь:** канд. техн. наук Калимулина Э.Ю. **Редактор:** канд. техн. наук Квинто Я.И.

Д-ра техн. наук: проф. Алескеров Ф.Т. (ГУ ВШЭ), проф. Алчинов А.И. (ИПУ РАН), проф. Андриевский Б.Р. (ИПМ РАН), проф. Афанасьев В.Н. (МИЭМ), проф. Бахтадзе Н.Н. (ИПУ РАН), проф. Бурков В.Н. (ИПУ РАН), проф. Вишневский В.М. (ИПУ РАН), Галаев А.А. (ИПУ РАН), д-р физ.-мат. наук проф. Ерешко Ф.И. (ВЦ РАН), д-ра техн. наук: Зоркальцев В.И. (ИСЭМ СО РАН), проф. Калашников А.О. (ИПУ РАН), проф. Калянов Г.Н. (ГУ ВШЭ), проф. Каравай М.Ф. (ИПУ РАН), д-р экон. наук, проф. Ключков В.В. (ИПУ РАН), д-р техн. наук Коргин Н.А. (ИПУ РАН), д-ра физ.-мат. наук: проф. Кушнер А.Г., проф. Лазарев А.А. (МФТИ), д-ра техн. наук: проф. Лебедев В.Г. (ИПУ РАН), проф. Мандель А.С. (ИПУ РАН), д-р биол. наук проф. Михальский А.И., д-р физ.-мат. наук, проф. Непейвода Н.Н. (ИПС РАН), д-р экон. наук, проф. Нижегородцев Р.М. (ИПУ РАН), д-р техн. наук, проф. Орлов А.И. (МГТУ), д-ра физ.-мат. наук: проф. Рапопорт Л.Б. (ИПУ РАН), проф. Райгородский А.М. (МГУ), проф. Савватеев А.В. (РЭШ), д-ра техн. наук: проф. Самуйлов К.Е. (РУДН), проф. Сидельников Ю.В. (МАИ), Совлуков А.С. (ИПУ РАН), д-ра физ.-мат. наук: проф. Соловьев С.Ю. (МГУ), проф. Угольницкий Г.А. (ЮФУ), проф. Уткин В.А. (ИПУ РАН), проф. Хоботов Е.Н. (МГТУ), д-ра физ.-мат. наук: доцент Чеботарев П.Ю. (ИПУ РАН), проф. Чхартишвили А.Г. (ИПУ РАН), проф. Щербаков П.С. (ИПУ РАН).

### РЕГИОНАЛЬНЫЕ РЕДАКЦИОННЫЕ СОВЕТЫ

**Арзамас** – д-р физ.-мат. наук проф. Пакшин П.В. **Волгоград** – д-ра физ.-мат. наук: проф. Воронин А.А., проф. Лосев А.Г. (ВолГУ); **Воронеж** – д-р техн. наук, проф. Баркалов С.А., д-р физ.-мат. наук, проф. Головинский П.А. (ВГАСУ), д-р техн. наук, проф. Подвальный С.Л. (ВГТУ); **Иркутск** – академик РАН Бычков И.В., д-р физ.-мат. наук, проф. Лакеев А.В. (ИДСТУ СО РАН); **Казань** – д-р физ.-мат. наук, проф. Маликов А.И., д-р техн. наук, проф. Сиразетдинов Р.Т. (КГТУ-КАИ); **Липецк** – д-ра техн. наук: проф. Погодаев А.К., Сараев П.В. (ЛГТУ); **Самара** – д-ра экон. наук: проф. Богатырев В.Д., проф. Гераськин М.И., д-р техн. наук, проф. Засканов В.Г. (СГАУ); **Петрозаводск** – д-р физ.-мат. наук, проф. Мазалов В.В., д-р техн. наук, доц. Печников А.А. (ИПМИ КарНЦ РАН); **Санкт-Петербург** – д-р физ.-мат. наук: проф. Петросян Л.А. (СПбГУ), д-р техн. наук проф. Фуртат И.Б. (ИПМ РАН); **Старый Оскол** – д-р техн. наук, проф. Еременко Ю.И. (СТИ).

**Адрес редакции:** 117997, г. Москва, ул. Профсоюзная, д. 65.

**Адрес в интернете:** [ubs.mtas.ru](http://ubs.mtas.ru).

## СОДЕРЖАНИЕ

### *Системный анализ*

---

**Зверкина Г.А.**

*Об экспоненциальной сходимости коэффициента готовности.....* 5

**Щеголев А.А.**

*Об оценках скорости сходимости однородных нелинейных цепей Маркова в дискретном времени* 36

### *Сетевые модели в управлении*

---

**Чаплинская Н.В.**

*Исследование полных однородных ресурсных сетей с «жадными» вершинами: зона «достаточного большого» ресурса .....* 49

### *Управление в социально-экономических системах*

---

**Шевыренков М.Ю.**

*К вопросу о корректировке постановки управленческой задачи* 67

### *Управление техническими системами и технологическими процессами*

---

**Клочков В.В., Топоров Н.Б., Егошин С.Ф.**

*Интегрированные авиационные системы* 94

### *Программы и системы моделирования объектов, средств и систем управления*

---

**Блинова О.В., Васьковский С.В., Рыков Я.В.**

*Проектирование сетей связи быстрого развертывания с использованием программного приложения анализа конфигурации сети* 121

## ОБ ЭКСПОНЕНЦИАЛЬНОЙ СХОДИМОСТИ КОЭФФИЦИЕНТА ГОТОВНОСТИ<sup>1</sup>

Зверкина Г. А.<sup>2</sup>

(ФГБУН Институт проблем управления  
им. В.А. Трапезникова РАН, Москва,

ФГАОУВО «Российский университет транспорта», Москва)

*В теории надёжности важную роль играет знание значения (или оценок значения) коэффициента готовности, т.е. вероятности того, что в заданный момент времени устройство исправно. В большинстве случаев вычисление стационарного значения коэффициента готовности не вызывает больших сложностей. Однако в реальных приложениях надо знать, как скоро значение коэффициента готовности станет достаточно близким к своему предельному значению, т.е. важно знать скорость сходимости коэффициента готовности к его стационарному значению. В тех случаях, когда распределения времён работы и восстановления являются экспоненциальными, вопрос о скорости сходимости коэффициента готовности – это вопрос о скорости сходимости решения уравнений Колмогорова – Чэпмена с постоянными коэффициентами. Такая задача разрешима с помощью операционного исчисления. Также заметим, что большинство процессов, описывающих поведение систем надёжности, являются регенерирующими, и для них известен тип скорости сходимости распределения к стационарному – экспоненциальное или степенное. Строгие оценки скорости сходимости для регенерирующих процессов могут быть получены с помощью неравенства Лордена, когда известно распределение длины одного периода регенерации. Но для систем надёжности можно учитывать и конкретные особенности исследуемого регенерирующего периода. В настоящей статье показан метод построения строгой экспоненциальной оценки скорости сходимости коэффициента готовности для одного восстанавливаемого элемента.*

**Ключевые слова:** коэффициент готовности, восстанавливаемый элемент, оценка скорости сходимости, экспоненциальная сходимость.

---

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ, грант №20-01-00575А.

Автор признателен Л.Г. Афанасьевой за ценное обсуждение и конструктивную критику статьи, а также анонимному рецензенту за указание на погрешности в оформлении текста.

<sup>2</sup> Галина Александровна Зверкина, к.ф.-м.н., доцент (zverkina@gmail.com).

## 1. Введение

Как известно, коэффициент готовности  $k_{\Gamma}(t)$  – это вероятность того, что в момент времени  $t$  изучаемый элемент (устройство, система) находится в рабочем состоянии. Предполагается, что этот элемент (устройство, система) может находиться в двух состояниях: исправном (обозначаемом  $\{1\}$ ) и неисправном (обозначаемом  $\{2\}$ ). Этим состояниям соответствуют последовательности случайных периодов времени  $\xi_k^{(1)}, \xi_k^{(2)}$ , где  $\xi_k^{(1)}$  –  $k$ -й период безотказной работы элемента, а  $\xi_k^{(2)}$  –  $k$ -й период его восстановления (замены). Периоды рабочего состояния элемента  $\xi_k^{(1)}$  – случайные величины с функцией распределения  $F^{(1)}(s) = \mathbf{P} \left\{ \xi_k^{(1)} \leq s \right\} = \mathbf{P} \left\{ \xi^{(1)} \leq s \right\}$ , периоды его неисправного состояния (время восстановления)  $\xi_k^{(2)}$  – случайные величины с функцией распределения  $F^{(2)}(s) = \mathbf{P} \left\{ \xi_k^{(2)} \leq s \right\} = \mathbf{P} \left\{ \xi^{(2)} \leq s \right\}$ ; случайные величины  $\xi_k^{(j)}, \xi_m^{(i)}$  независимы в совокупности (здесь и везде далее  $i, j = \overline{1, 2}$ ).

Следуя [2, Гл. 2], дадим определение.

Определение 1. Пусть некий элемент, начав работу в момент  $t = 0$  и проработав случайное время  $\xi_1^{(1)}$ , выходит из строя, затем он восстанавливается (заменяется исправным) в течение случайного времени  $\xi_1^{(2)}$ ; восстановленный элемент работает время  $\xi_2^{(1)}$  и затем восстанавливается время  $\xi_2^{(2)}$ ; и т.д. Положим

$$t_k \stackrel{\text{def}}{=} \sum_{m=1}^k \left( \xi_m^{(1)} + \xi_m^{(2)} \right);$$

$$t'_k \stackrel{\text{def}}{=} \sum_{m=0}^{k-1} \left( \xi_m^{(1)} + \xi_m^{(2)} \right) + \xi_k^{(1)} = t_k + \xi_k^{(1)};$$

моменты  $t_k$  есть моменты восстановления элемента, а моменты  $t'_k$  – моменты его отказов; при этом  $\xi_k^{(1)} = t'_k - t_k$ ,  $\xi_k^{(2)} = t_k - t'_{k-1}$ .

Коэффициент готовности  $k_{\Gamma}(t)$  элемента в момент  $t$  определяется как вероятность того, что в момент  $t$  элемент исправен:

$$k_{\Gamma}(t) \stackrel{\text{def}}{=} \mathbf{P} \left( \bigcup_{k=1}^{\infty} \{t \in (t_k; t'_k)\} \right). \triangleright$$

Случайный процесс, описывающий поведение такого восстанавливаемого элемента, – регенерирующий. Из классической теории восстановления (см., например, [3, Гл. 2] и [5, Ч. I, Гл. 5]) следует, что в случае, когда  $\xi_k^{(1)}$  и  $\xi_k^{(2)}$  имеют *непериодическое распределение* и конечные математические ожидания

$$T_i \stackrel{\text{def}}{=} \mathbf{E} \xi^{(i)} < \infty,$$

существует предельное (стационарное) значение коэффициента готовности:  $\lim_{t \rightarrow \infty} k_{\Gamma}(t) = k_{\Gamma} = \frac{T_1}{T_1 + T_2}$ .

Наиболее просто исследуется ситуация, когда  $F^{(1)}(s)$  и  $F^{(2)}(s)$  – экспоненциальные распределения с параметрами  $\lambda_1$  и  $\lambda_2$  соответственно. В этом случае  $\lim_{t \rightarrow \infty} k_{\Gamma}(t) = k_{\Gamma} = \frac{\lambda_2}{\lambda_1 + \lambda_2}$ , а скорость сходимости  $k_{\Gamma}(t)$  к  $k_{\Gamma}$  экспоненциальна; имеет место точная оценка:

$$(1) \quad 0 < k_{\Gamma}(t) - k_{\Gamma} = \frac{\lambda_2}{\lambda_1 + \lambda_2} e^{-(\lambda_1 + \lambda_2)t},$$

см. [2, Гл. 2, § 3], [5, Ч. I, Гл. 5].

Из теоремы 3 [1, Приложение 1 (Стр. 330)] (или теоремы 10.7.4 [10, С. 417]) следует, что если при некотором  $c > 0$  выполнено  $\mathbf{E} e^{c(\xi_k^{(1)} + \xi_k^{(2)})} < \infty$ , то  $\forall \alpha \in (0; c) \exists K(\alpha)$  такое, что  $\forall t \geq 0$  верно

$$(2) \quad |k_{\Gamma}(t) - k_{\Gamma}| < K(\alpha) e^{-\alpha t},$$

однако указанные результаты не дают каких-либо оценок для коэффициента скорости сходимости  $K(\alpha)$ .

Если при  $\alpha > c$  выполнено  $\mathbf{E} e^{c(\xi_k^{(1)} + \xi_k^{(2)})} = \infty$ , то  $\lim_{\alpha \uparrow c} K(\alpha) = \infty$ , т.е. при малых  $c - \alpha$  применение неравенства (2)

приводит к очень большим значениям (неизвестной) постоянной  $K(\alpha)$  (естественно, если известна величина этой постоянной  $K(\alpha)$  или хотя бы её оценка).

Целью настоящей работы является получение оценки для  $K(\alpha)$  в случае, когда время безотказной работы элемента задаётся ограниченной отделённой от нуля *интенсивностью*  $\lambda(\cdot)$ , а время его восстановления имеет конечный экспоненциальный момент.

Во второй части работы сформулирован основной результат и изложена идея его доказательства, в третьей части приведены примеры, в четвёртой части приводятся вспомогательные факты, и в пятой и шестой частях – доказательство основного результата.

## 2. Обозначения и основной результат

Итак, имеется элемент, который может находиться в двух состояниях – рабочем  $\{1\}$  и нерабочем  $\{2\}$ . Времена пребывания в этих состояниях случайны и независимы, их распределения имеют соответственно функции распределения  $F^{(1)}(s)$  и  $F^{(2)}(s)$ , при этом *интенсивность* восстановления элемента зависит от времени, прошедшего с момента отказа. Это значит, что если в момент  $t$  элемент находится в состоянии  $\{1\}$  и попал в это состояние  $s$  единиц времени назад, то вероятность того, что на интервале  $(t; t + \Delta t)$  он изменит состояние ( $\{1\} \rightarrow \{2\}$ ), равна  $\lambda(s)\Delta t + o(\Delta t)$ . Распределение  $\xi^{(1)}$  – времени исправного состояния элемента – определяется *интенсивностью*  $\lambda(s)$ , т.е. функцией, связанной с ф.р. времени безотказной работы  $F^{(1)}(s)$ :

$$(3) \quad \mathbf{P} \{ \xi^{(1)} \leq s \} = F^{(1)}(s) = 1 - \exp \left( - \int_0^s \lambda(u) du \right);$$

$$f^{(1)}(s) \stackrel{\text{def}}{=} (F^{(1)}(s))' = \lambda(s) \exp \left( - \int_0^s \lambda(u) du \right).$$

Таким образом, положительная сл.в. с абсолютно непрерывной ф.р. может быть определена как с помощью ф.р. и плотности распределения, так и с помощью интенсивности.



Для того чтобы определённая в (3) функция  $F^{(1)}(s)$  была функцией распределения, достаточно, чтобы  $\lambda(s)$  была неотрицательна, измерима по Борелю и  $\int_0^{\infty} \lambda(s) ds = \infty$ .

Распределение  $\xi^{(2)}$  – времени восстановления элемента – имеет функцию распределения  $F^{(2)}(s)$ . Везде в дальнейшем  $\xi_k^{(i)} \stackrel{\mathcal{D}}{=} \xi^{(i)}$ .

Поведение восстанавливаемого элемента описывается альтернирующим процессом восстановления (см. [5]):

$$n_t = \begin{cases} 1, & \text{если элемент в момент } t \text{ исправен;} \\ 2, & \text{если элемент в момент } t \text{ неисправен.} \end{cases}$$

Дополним этот процесс  $n_t$  ещё одной компонентой – временем, прошедшим с момента попадания процесса  $n_t$  в то состояние, в котором он находится в момент  $t$ , и получим *марковский* случайный процесс  $X_t = (n_t, x_t)$ , определённый так:

$$X_t \stackrel{\text{def}}{=} \begin{cases} (1, x), & \text{если элемент исправен и находится} \\ & \text{в исправном состоянии } x \text{ единиц времени;} \\ (2, x), & \text{если элемент неисправен и находится} \\ & \text{в неисправном состоянии } x \text{ единиц времени.} \end{cases}$$

Пространство состояний процесса  $X_t$  есть  $\mathcal{X} = \{1, 2\} \times \mathbb{R}_+ = \mathcal{S}_1 \cup \mathcal{S}_2$ , где  $\mathcal{S}_1 \stackrel{\text{def}}{=} \{\{1\} \times \mathbb{R}_+\}$  (элемент исправен) и  $\mathcal{S}_2 \stackrel{\text{def}}{=} \{\{2\} \times \mathbb{R}_+\}$  (элемент неисправен);  $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$ .

В момент времени  $t$  процесс  $X_t$  имеет на пространстве  $\mathcal{X}$  распределение  $\mathcal{P}_t$ , т.е.  $\mathbf{P}\{X_t \in A\} = \mathcal{P}_t(A)$  для  $A \in \mathcal{B}(\mathcal{X})$ . Обозначим  $n(X_t) \stackrel{\text{def}}{=} n_t$ ,  $x(X_t) \stackrel{\text{def}}{=} x_t$ .

В дальнейшем под состоянием процесса  $X_t$  понимаем пару  $(n_t, x_t)$ .

Обычно предполагают, что восстанавливаемый элемент начинает свою работу в исправном состоянии в начальный момент времени  $t = 0$ :  $X_0 = (1, 0)$ . Мы будем предполагать, что в начальный момент времени  $t = 0$  процесс  $X_t$  может принимать

любое значение из  $\mathcal{X}$ :  $X_0 = (n_0, x_0)$ . Поэтому распределение процесса  $X_t$  зависит от начального состояния: будем обозначать это распределение  $\mathcal{P}_t^{X_0}$  или  $\mathcal{P}_t^{(n_0, x_0)}$ . Коэффициент готовности элемента с начальным состоянием  $X_0$  определяется как

$$k_{\Gamma}^{X_0}(t) \stackrel{\text{def}}{=} \mathbf{P}\{n(X_t) = 1\} = \mathbf{P}\{X_t \in \mathcal{S}_1\} = \mathcal{P}_t^{X_0}(\mathcal{S}_1).$$

## 2.1. ОБОЗНАЧЕНИЯ

2.1.1.  $\mathcal{T}(X_0) \stackrel{\text{def}}{=} \inf\{t > 0 : X_t = (1, 0)\}$  – первый момент регенерации указанного ниже типа процесса  $X_t$  (см. ниже раздел 5.1); вообще говоря, могут рассматриваться и другие точки регенерации, но для простоты изложения мы выбрали эти.

2.1.2.  $\xi_x^{(i)}$  – сл.в. с ф.р.  $F_x^{(i)}(s)$ , и

$$(4) \quad F_x^{(i)}(s) \stackrel{\text{def}}{=} 1 - \frac{1 - F^{(i)}(x + s)}{1 - F^{(i)}(x)}; \quad f_x^{(i)}(s) \stackrel{\text{def}}{=} \left(F_x^{(i)}(s)\right)'$$

$$\text{Если } F^{(2)}(x) = 1, \text{ то } \frac{1 - F^{(2)}(x + s)}{1 - F^{(2)}(x)} \stackrel{\text{def}}{=} 0; F_0^{(i)}(s) = F^{(i)}(s)$$

при  $F^{(i)}(0+) = 0$ .

При  $X_t \in \mathcal{S}_i$  сл.в.  $\xi_x^{(i)} \stackrel{\text{def}}{=} (\inf\{s > t : X_t \notin \mathcal{S}_i\} - t)$  – остаточное время пребывания процесса  $X_t$  в множестве  $\mathcal{S}_i$  – имеет ф.р.  $F_x^{(i)}(s)$ .

**Замечание 1.** Из (3) следует: найдутся такие положительные  $c$  и  $C$ , что при  $s \in \mathbb{R}_+$

$$f^{(1)}(s), f_x^{(1)}(s) \in (ce^{-Cs}; Ce^{-cs}); \\ F^{(1)}(s), F_x^{(1)}(s) \in (1 - e^{-cs}, 1 - e^{-Cs}),$$

поэтому для всех  $\alpha < c$

$$(5) \quad \mathbf{E} e^{\alpha \xi^{(1)}} \leq \frac{C}{c - \alpha}, \quad \mathbf{E} e^{\alpha \xi_x^{(1)}} \leq \frac{C}{c - \alpha}.$$

Из этого следует, что сумма периодов работы и восстановления имеет конечное математическое ожидание, а, значит, распределение  $\mathcal{P}_t^{X_0}$  состояния рассматриваемой восстанавливаемой системы надёжности при любом начальном состоянии  $X_0$  процесса  $X_t$  слабо сходится к стационарному предельному инвариантному вероятностному распределению  $\mathcal{P}$  на  $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ .  $\triangleright$

$$2.1.3. T_1 \stackrel{\text{def}}{=} \mathbf{E} \xi^{(1)}; \quad T_2 \stackrel{\text{def}}{=} \mathbf{E} \xi^{(2)}; \quad k_{\Gamma} \stackrel{\text{def}}{=} \frac{T_1}{T_1 + T_2}.$$

2.1.4.  $\tilde{\xi}^{(i)}$  – сл.в. с плотностью распределения

$$(6) \quad \tilde{f}^{(i)}(s) = \frac{1 - F^{(i)}(s)}{\mathbf{E} \xi^{(i)}} = \frac{1 - F^{(i)}(s)}{T_i}.$$

$$2.1.5. \quad \Theta \stackrel{\text{def}}{=} \frac{\mathbf{E} (\xi^{(2)})^2}{2T_2};$$

$$p = p(c, C, R, \Theta) \stackrel{\text{def}}{=} \frac{ce^{-C\Theta}}{C} \left(1 - \frac{\Theta}{R}\right); \quad q \stackrel{\text{def}}{=} 1 - p;$$

$$R^* = R^*(C, \Theta) \stackrel{\text{def}}{=} \frac{\Theta + \sqrt{\Theta^2 + 4\frac{\Theta}{C}}}{C}.$$

$$2.1.6. \quad \widehat{M}_\alpha \stackrel{\text{def}}{=} \int_0^\infty \int_0^\infty e^{\alpha s} (1 - F^{(2)}(s + u)) \, ds \, du.$$

**Замечание 2.** Для конечности  $\widehat{M}_\alpha$  достаточно конечности  $\sup_{x \geq 0} (1 - F^{(2)}(s)) e^{-cs} < \infty$  при некотором  $c > \alpha$ .  $\triangleright$

$$2.1.7. \quad M_\alpha(n, x) \stackrel{\text{def}}{=} \begin{cases} \mathbf{E} e^{\alpha \xi^{(2)}} \mathbf{E} e^{\alpha \xi_x^{(1)}} & \text{при } n = 1; \\ \mathbf{E} e^{\alpha \xi_x^{(2)}} & \text{при } n = 2. \end{cases}$$

$$2.1.8. \quad \overline{M}_\alpha \stackrel{\text{def}}{=} \frac{1}{T_1 + T_2} \left( \frac{\alpha C \mathbf{E} e^{\alpha \xi^{(2)}}}{(c - \alpha)c} + \alpha \widehat{M}_\alpha^{(2)} + T_1 \mathbf{E} e^{\alpha \xi^{(2)}} + T_2 \right).$$

$$2.1.9. \quad \mathfrak{K}(\alpha) = \mathfrak{K}(\alpha, c, C, X_0) \stackrel{\text{def}}{=} \frac{C e^{\alpha R^*} M_\alpha(X_0) \overline{M}_\alpha}{(c - \alpha) (1 - \mathbf{E} e^{\alpha \xi^{(1)}} \mathbf{E} e^{\alpha \xi^{(2)}} q)}.$$

## 2.2. ОСНОВНОЙ РЕЗУЛЬТАТ

**Теорема 1.** Если выполнены условия

$$(7) \quad \begin{aligned} F^{(1)}(s) &= 1 - \exp\left(-\int_0^s \lambda(u) du\right); \\ \forall x \geq 0 \quad 0 < c < \lambda(x) \leq C < \infty, \\ \widehat{M}_c^{(2)} < \infty; \quad F^{(2)}(0+) &= 0, \\ \mathbf{E} e^{\alpha \xi^{(1)}} \mathbf{E} e^{\alpha \xi^{(2)}} q < 1 - \text{см. п. 2.1.5,} \end{aligned}$$

$u \forall t > 0$

$$(8) \quad \|\mathcal{P}_t^{X_0} - \mathcal{P}\|_{TV} \stackrel{\text{def}}{=} \stackrel{\text{def}}{\sup_{A \in \mathcal{B}(\mathcal{X})}} |\mathcal{P}_t^{X_0}(A) - \mathcal{P}(A)| \leq \mathfrak{K}(\alpha) e^{-\alpha t},$$

где постоянная  $\mathfrak{K}(\alpha)$  определена в п. 2.1.9. •

**Следствие 1.** В условиях теоремы 1

$$\left| k_{\Gamma}^{X_0}(t) - k_{\Gamma} \right| \leq \mathfrak{K}(\alpha) e^{-\alpha t}. \quad \triangleright$$

Следствие 1 вытекает из того, что

$$k_{\Gamma}^{X_0}(t) = \mathbf{P}\{n(X_t) = 1\} = \mathbf{P}\{X_t \in \mathcal{S}_1\} = \mathcal{P}_t^{X_0}(\mathcal{S}_1).$$

## 3. Идея доказательства теоремы 1

Доказательство теоремы основано на использовании *метода склеивания (coupling method)* – см. [8, 10]) марковских процессов.

Метод склеивания основан на оценке момента времени, когда происходит совпадение двух *независимых* марковских процессов с одинаковыми переходными функциями и произвольными начальными состояниями.

То есть мы рассматриваем две *независимые* версии процесса  $X_t - X'_t$  и  $X''_t$ , начинающиеся соответственно из начальных состояний  $X'_0$  и  $X''_0$ . В общем случае эти процессы могут совпасть (склеиться) только в момент перехода в одно и то же состояние (т.е. переход вида  $\{1\} \rightarrow \{2\}$  или  $\{2\} \rightarrow \{1\}$ ). Вероятность того, что эти процессы совпадут (при  $X'_0 \neq X''_0$ ), может оказаться

отличной от нуля только в том случае, когда оба процесса начинаются из состояния  $\{2\}$  и ф.р.  $F_2(s)$  имеет дискретную составляющую (напомним, что ф.р.  $F_1(s)$  абсолютно непрерывна).

Поэтому на некотором специально выбранном вероятностном пространстве строится пара процессов  $(Z'_t, Z''_t)$  такая, что маргинальные распределения компонент этого «парного» процесса совпадают с распределениями процессов  $X'_t, X''_t$  (в частности,  $Z'_0 = X'_0, Z''_0 = X''_0$ ), и при этом для всех  $t > T_{Z'_0, Z''_0} \stackrel{\text{def}}{=} \inf\{t > 0 : Z'_t = Z''_t\}$  выполнено равенство  $Z'_t = Z''_t$ . Следуя [7], эту пару называем *склейкой* процессов  $X'_t$  и  $X''_t$ .

Если сл.в.  $T_{Z'_0, Z''_0}$  конечна с вероятностью 1, то *склейка* называется *успешной*.

Наша цель – построить *успешную склейку*  $(Z'_t, Z''_t)$  процессов  $X'_t$  и  $X''_t$ , такую, что для некоторого  $\alpha > 0$  и

$$K(\alpha, Z'_0, Z''_0) = K(\alpha, c, C, Z'_0, Z''_0)$$

выполнено

$$\mathbf{E} e^{\alpha T_{Z'_0, Z''_0}} \leq K(\alpha, Z'_0, Z''_0) < \infty,$$

где  $c$  и  $C$  – постоянные из замечания 1.

В этом случае можно применить к процессам  $X'_t$  и  $X''_t$  основное неравенство склеивания (basic coupling inequality – см. ниже п. 6.3.7).

Пусть  $\tilde{X}_t$  – процесс, начальное *распределение* которого –  $\mathcal{P}$ ; это – стационарный процесс.

Если  $\int_{\mathcal{X}} K(\alpha, X'_0, X''_0) \mathcal{P}(dX''_0) \leq \mathfrak{K}(\alpha) = \mathfrak{K}(\alpha, c, C, X'_0)$ , то

$$\begin{aligned} \sup_{A \in \mathcal{B}(\mathcal{X})} \left| \mathbf{P}\{X'_t \in A\} - \mathbf{P}\{\tilde{X}_t \in A\} \right| &= \\ &= \sup_{A \in \mathcal{B}(\mathcal{X})} \left| \mathcal{P}_t^{X'_0}(A) - \mathcal{P}(A) \right| \leq \mathfrak{K}(\alpha, X'_0) e^{-\alpha t}, \end{aligned}$$

откуда и следует утверждение теоремы 1.

Полное доказательство теоремы 1 дано в разделах 5–6.

#### 4. Примеры использования полученных результатов

Для двух вариантов ф.р.  $F^{(1)}(s)$  и  $F^{(2)}(s)$  было проведено имитационное моделирование (10 000 000 реализаций, вычисления с 16 знаками после нуля) поведения процесса  $X_t$  с различными начальными условиями, вычислены эмпирические значения коэффициента готовности  $k_{\Gamma}^{эмн}(t)$  для  $t = 1, 2, \dots, 650$ , получена (несколько заниженная) эмпирическая оценка  $\mathfrak{K}^{эмн}(\alpha)$  величины  $\mathfrak{K}(\alpha)$ , которая была сравнена с  $\mathfrak{K}(\alpha) = \mathfrak{K}(\alpha, c, C, X_0)$ .

Причин того, что эмпирическая оценка  $\mathfrak{K}^{эмн}(\alpha)$  занижена, две.

Во-первых, при  $\Delta^{эмн}(t) \stackrel{\text{def}}{=} |k_{\Gamma}^{эмн}(t) - k_{\Gamma}| < 0,0004$  разброс  $\Delta^{эмн}(t)$  сравним с величиной  $\Delta^{эмн}(t)$ . Это следствие накопления ошибок округления при вычислении значений сл.в.  $\xi^{(i)}$ ; увеличение числа реализаций здесь не даёт улучшения точности (см. рис.1 и таблицу 1).

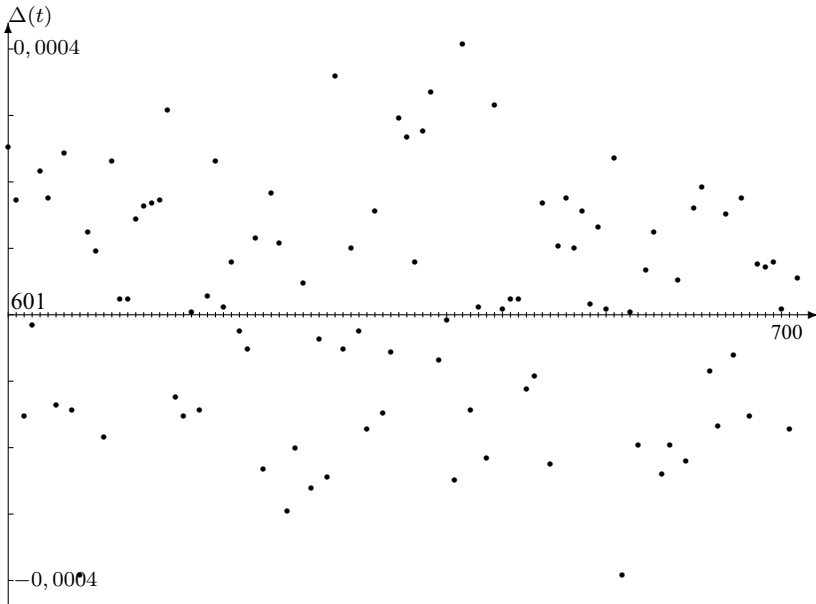


Рис. 1.  $\Delta(t)$  для  $t = 601, 602, \dots, 700$

Поэтому вычисление  $\mathfrak{K}^{mn}(\alpha)$  проводилось только по нескольким из эмпирических данных – по значениям  $\Delta(t)$  для первых 10–15 единиц времени.

Таблица 1.  $\Delta(t)$  для примера 1 (начальное условие  $X_0 = (2, 0)$ )

$t$	$\Delta(t)$	$t$	$\Delta(t)$
0	-0,5	16	0,0000432
1	0	17	0,0002705
2	-0,1322323	18	0,0001697
3	0,0032313	19	-0,0000246
4	0,0045013	20	0,0000847
5	-0,0005616	21	-0,0000148
6	-0,0003203	22	-0,0002263
7	0,0002989	23	-0,0000415
8	-0,0000282	24	-0,0000685
9	0,0002417	25	0,0004427
10	-0,0001726	26	-0,0003706
11	-0,0000200	27	-0,0003147
12	0,0001399	28	-0,0000076
13	-0,0001912	29	-0,0001412
14	0,0002538	30	-0,0000692
15	-0,0000148	31	-0,0001873

Также было проведено сравнение величин  $\mathfrak{T}_\varepsilon^{mn}(X_0) \stackrel{\text{def}}{=} \inf\{t > 0 : \forall s > t \quad |k_\Gamma^{mn}(s) - k_\Gamma| < \varepsilon\}$ , полученных в численных экспериментах (здесь точность определения  $\mathfrak{T}_\varepsilon^{mn}(X_0)$  представляется удовлетворительной) и  $\mathfrak{T}_\varepsilon(X_0) \stackrel{\text{def}}{=} \inf\{t > 0 : \forall s > t \quad |k_\Gamma(s) - k_\Gamma| < \varepsilon\}$ , полученных из оценки (8).  $\mathfrak{T}_\varepsilon(X_0)$  – это время, по истечении которого можно с точностью  $\varepsilon$  определить  $k_\Gamma$  по результатам наблюдений за процессом  $X_t$ .

Поскольку при имитационном моделировании значение  $k_\Gamma(t)$  вычислялось в целочисленные моменты времени, для  $\mathfrak{T}_\varepsilon^{mn}(X_0)$  определялся интервал единичной длины, в который попало значение  $\mathfrak{T}_\varepsilon^{mn}(X_0)$  – см. таблицы 3 и 4.

4.1. ПРИМЕР 1

Пусть  $F^{(1)}(s) = 1 - e^{-s}$ ,  $\xi^{(2)} \equiv 1$  (время восстановления элемента постоянно). В этом случае  $C = c = 1$ ;  $\Theta = 1$ ;  $R^* \approx 1,618$ ;<sup>1</sup>  $q \approx 0,9243$ ;  $\mathbf{E} e^{\xi^{(1)}} = \frac{1}{1 - \alpha}$  при  $\alpha < 1$ ;  $\mathbf{E} e^{\alpha \xi^{(1)}} \mathbf{E} e^{\alpha \xi^{(2)}} q < 1$  при  $\alpha < 0,039$ , т.е. формула (8) применима для  $\alpha \in (0; 0,03899)$ .

Поскольку для любых  $a > 0$  верно равенство  $F_a^{(1)}(s) \equiv F_0^{(1)}(s)$ , распределения  $\mathcal{P}_t^{(1,0)}$  и  $\mathcal{P}_t^{(1,\cdot)}$  совпадают.

Заметим, что при любом начальном значении  $X_0$  максимум  $|\Delta(t)|$  достигается при  $t = 0$ :  $|\Delta(0)| = 0,5$ , поскольку  $\mathbf{P}\{X_0 \in \mathcal{S}_1\} = 0$  или  $\mathbf{P}\{X_0 \in \mathcal{S}_1\} = 1$ , а при  $k \in N$  величина  $|\Delta(t)|$  убывает очень быстро. То есть при  $\alpha \in (0; 0,03899)$  величина оценки  $K^{\varepsilon mn}(\alpha)$  не зависит от  $\alpha$ :  $K^{\varepsilon mn}(\alpha) \equiv 0,5$ .

Поэтому также вычислялись оценки для  $\tilde{\mathfrak{K}}^{\varepsilon mn}$  без учёта значений  $k_{\Gamma}^{\varepsilon mn}(t)$  для  $t \in [0; 1)$ .

Оценки  $\mathfrak{K}(\alpha)$  приведены в таблице 3; поскольку диапазон изменения  $\alpha$  очень мал ( $\alpha < 0,03899$ ), отношения величин  $M_\alpha(X_0)$  для рассмотренных начальных значений  $X_0$  равны примерно 1, и поэтому отношения величин  $\mathfrak{K}(\alpha, X_0)$  также близки к 1 (см. п. 2.1.9).

<sup>1</sup> Золотое сечение.

Таблица 2. Пример 1:  $\mathfrak{T}_\varepsilon^{\varepsilon mn}$ ,  $\mathfrak{T}_\varepsilon$ ; минимальные значения  $\mathfrak{T}_\varepsilon$  достигаются при  $\alpha \approx 0,0341$ ,  $\alpha \approx 0,0353$  и  $\alpha \approx 0,0360$  соответственно

	$\varepsilon = 0,1$	$\varepsilon = 0,01$	$\varepsilon = 0,001$
$\mathfrak{T}_\varepsilon^{\varepsilon mn}(2,0)$	(2;3)	(2;3)	(5;6)
$\mathfrak{T}_\varepsilon(2,0)$	90,213	118,993	147,025
$\mathfrak{T}_\varepsilon^{\varepsilon mn}(2, \frac{1}{2})$	(2;3)	(2;3)	(4;5)
$\mathfrak{T}_\varepsilon(2, \frac{1}{2})$	89,996	118,776	146,808
$\mathfrak{T}_\varepsilon^{\varepsilon mn}(1, \cdot)$	(0;1)	(0;1)	(4;5)
$\mathfrak{T}_\varepsilon(1, \cdot)$	90,66	119,435	147,468



Таблица 3. Пример 1: значения  $\mathfrak{K}(\alpha)$

$X_0$	$\alpha$	0	0,01	0,02	0,03	0,03899
(2,0)	$\mathfrak{K}^{ЭМП}$	0,5	0,5	0,5	0,5	0,5
	$\tilde{\mathfrak{K}}^{ЭМП}$	0,132	0,135	0,138	0,14	0,143
	$\mathfrak{K}$	13,203	18,414	29,154	63,873	$1,787 \cdot 10^6$
$(2, \frac{1}{2})$	$\mathfrak{K}^{ЭМП}$	0,5	0,5	0,5	0,5	0,5
	$\tilde{\mathfrak{K}}^{ЭМП}$	0,132	0,135	0,138	0,14	0,143
	$\mathfrak{K}$	13,203	18,322	28,864	62,922	$1,752 \cdot 10^6$
(1,·)	$\mathfrak{K}^{ЭМП}$	$3,36 \cdot 10^{-3}$	$3,4 \cdot 10^{-3}$	$3,43 \cdot 10^{-3}$	$3,47 \cdot 10^{-3}$	$3,5 \cdot 10^{-3}$
	$\tilde{\mathfrak{K}}^{ЭМП}$	$3,36 \cdot 10^{-3}$	$3,4 \cdot 10^{-3}$	$3,43 \cdot 10^{-3}$	$3,47 \cdot 10^{-3}$	$3,5 \cdot 10^{-3}$
	$\mathfrak{K}$	13,203	18,678	29,735	63,135	$2,532 \cdot 10^5$

#### 4.2. ПРИМЕР 2

Пусть  $F^{(1)}(s) = 1 - e^{-4s}$ ,  $F^{(2)}(s) = 1 - e^{-s^2}$ ; заметим, что в этом случае интенсивность восстановления элемента  $\frac{(F^{(2)}(s))'}{1 - F^{(2)}(s)} = 2s$ .

Далее,

$$\begin{aligned}
 c = C = 0,7; & \quad \mathbf{E} \xi^{(1)} \approx 1,428; & \quad \mathbf{E} (\xi^{(1)})^2 \approx 4,082; \\
 \mathbf{E} \xi^{(2)} = \frac{\sqrt{\pi}}{2} \approx 0,8862; & \quad \mathbf{E} (\xi^{(2)})^2 = 1; & \quad \Theta = \mathbf{E} \xi^{(1)} \approx 1,428; \\
 R^* \approx 2,3115; & \quad q \approx 0,9243; & \quad k_{\Gamma} \approx 0,6171;
 \end{aligned}$$

функция  $q\mathbf{E}e^{\alpha\xi^{(1)}}\mathbf{E}e^{\alpha\xi^{(2)}}$  равна 1 при  $\alpha = \alpha_0 \approx 0,04117$ , т.е. для всех  $\alpha < 0,04117$  применима оценка (8).

Сравнение эмпирических результатов и результата применения теоремы 1 приведено в таблице 4.

В этом случае  $\frac{\mathfrak{K}(\alpha, (2, 0))}{\mathfrak{K}(\alpha, (1, \cdot))} \in (0,95; 1)$ , но  $\frac{\mathfrak{K}(\alpha, (2, x))}{\mathfrak{K}(\alpha, (1, \cdot))} < 0,05$  при  $x \geq 10$ , поэтому оценки  $\mathfrak{T}_{\varepsilon}$  для  $X_0 = (2, 0)$  и  $X_0 = (1, \cdot)$  близки, а оценки  $\mathfrak{T}_{\varepsilon}$  для  $X_0 = (2, 10)$  и  $X_0 = (2, 25)$  существенно от них отличаются (таблица 5).

Таблица 4. Пример 2:  $\mathfrak{F}_\varepsilon^{\alpha mn}$ ,  $\mathfrak{F}_\varepsilon$ ; указаны значения  $\alpha$ , соответствующие указанным значениям  $\mathfrak{F}_\varepsilon$

	$X_0$	$\varepsilon = 0,1$	$\varepsilon = 0,01$	$\varepsilon = 0,001$
$\mathfrak{F}_\varepsilon$	(2,0)	86,008	113,263	139,809
$\mathfrak{F}_\varepsilon^{\alpha mn}$	(2,0)	(1;2)	(2;3)	(3;4)
$\alpha$	(2,0)	0,036	0,0373	0,038
$\mathfrak{F}_\varepsilon$	(2,10)	47,627	77,154	104,729
$\mathfrak{F}_\varepsilon^{\alpha mn}$	(2,10)	(0;1)	(1;2)	(2;3)
$\alpha$	(2,10)	0,0317	0,0354	0,037
$\mathfrak{F}_\varepsilon$	(2,25)	34,409	67,269	93,919
$\mathfrak{F}_\varepsilon^{\alpha mn}$	(2,25)	(0;1)	(0;1)	(2;3)
$\alpha$	(2,25)	0,0279	0,0373	0,0373
$\mathfrak{F}_\varepsilon$	(1,·)	86,45	113,705	140,252
$\mathfrak{F}_\varepsilon^{\alpha mn}$	(1,·)	(0;1)	(0;1)	(2;3)
$\alpha$	(1,·)	0,036	0,0373	0,038

### 4.3. ПРИМЕР 3

Пусть  $F^{(1)}(s) = 1 - e^{-\lambda s}$ ,  $F^{(2)}(s) = 1 - e^{-\mu s}$ ; в этом случае имеет место оценка (1);  $k_\Gamma = \frac{\mu}{\lambda + \mu}$ .

Оценим скорость сходимости  $k_\Gamma(t)$  к предельному значению, используя оценки из доказательства теоремы 1, но несколько улучшив точность с использованием свойств показательного распределения: входящие в  $\mathfrak{R}(\alpha)$  величины либо вычисляются точно, либо могут быть более точно оценены.

1.  $\mathbf{P}(\mathcal{E}_k)$  (см. 6.3.4) вычисляется явно:  $\mathbf{P}(\mathcal{E}_k) = \frac{\mu}{2(\lambda + \mu)}$ ; чтобы эта вероятность была больше, будем считать, что  $\mu > \lambda$ .

2. Из того, что

$$(9) \quad F_a^{(i)}(s) \equiv F^{(i)}(s),$$

следует, что  $\varkappa_{x,y} \equiv 1$ .

3. Поскольку  $\mathbf{P} \left\{ \xi_1^{(1)} > \xi_2^{(1)} + \xi_2^{(2)} \right\} = 1 - q = \frac{\mu}{2(\lambda + \mu)}$ ;

$$q = \frac{2\lambda + \mu}{2(\lambda + \mu)} \quad \left( \text{здесь } \hat{\xi}_k^{(i)} \stackrel{\mathcal{D}}{=} \xi^{(i)} \right).$$

Таблица 5. Пример 2: значения  $\mathfrak{K}(\alpha)$

$X_0$	$\alpha$	0	0,01	0,02	0,03	0,04117
(2,0)	$\mathfrak{K}^{\text{ЭМН}}(\alpha)$	0,6171	0,6171	0,6171	0,6171	0,6171
	$\tilde{\mathfrak{K}}^{\text{ЭМН}}(\alpha)$	0,119	0,12	0,121	0,122	0,124
	$\mathfrak{K}(\alpha)$	13,203	18,287	28,239	56,143	$2,428 \cdot 10^5$
(2,10)	$\mathfrak{K}^{\text{ЭМН}}(\alpha)$	0,6171	0,6171	0,6171	0,6171	0,6171
	$\tilde{\mathfrak{K}}^{\text{ЭМН}}(\alpha)$	0,008	$8,414 \cdot 10^{-3}$	$8,498 \cdot 10^{-3}$	$8,584 \cdot 10^{-3}$	$8,68 \cdot 10^{-3}$
	$\mathfrak{K}(\alpha)$	0,657	0,902	1,382	2,724	$1,167 \cdot 10^4$
(2,25)	$\mathfrak{K}^{\text{ЭМН}}(\alpha)$	0,6171	0,6171	0,6171	0,6171	0,6171
	$\tilde{\mathfrak{K}}^{\text{ЭМН}}(\alpha)$	0,005	$5,189 \cdot 10^{-3}$	$5,241 \cdot 10^{-3}$	$5,294 \cdot 10^{-3}$	$5,354 \cdot 10^{-3}$
	$\mathfrak{K}(\alpha)$	0,264	0,362	0,554	1,092	$4,677 \cdot 10^4$
(1,.)	$\mathfrak{K}^{\text{ЭМН}}(\alpha)$	0,3829	0,3829	0,3829	0,3829	0,3829
	$\tilde{\mathfrak{K}}^{\text{ЭМН}}(\alpha)$	0,003	$3,397 \cdot 10^{-3}$	$3,431 \cdot 10^{-3}$	$3,465 \cdot 10^{-3}$	$3,504 \cdot 10^{-3}$
	$\mathfrak{K}(\alpha)$	13,203	18,471	28,815	57,879	$2,532 \cdot 10^5$

4. Вместо  $M_\alpha(n, x)$  будем использовать

$$\tilde{M}_\alpha(n, x) = \begin{cases} 1 & \text{при } n = 1; \\ \frac{\mu}{\mu - \alpha} & \text{при } n = 2. \end{cases}$$

5. Из (9) следует, что в качестве  $\mathcal{T}_0$  (14) для процесса  $Z_t$  можно взять момент  $\tilde{\mathcal{T}}_0(Z_0) \stackrel{\text{def}}{=} \inf\{t \geq 0 : n(Z_t) = 1\}$ ; соответственно

$$\tilde{\mathcal{T}}_0(n, x) \stackrel{\mathcal{D}}{=} \begin{cases} 0 & \text{при } n = 1; \\ \xi_x^{(2)} \stackrel{\mathcal{D}}{=} \xi^{(2)} & \text{при } n = 2, \end{cases}$$

поэтому  $\mathbf{E} e^{\alpha \tilde{\mathcal{T}}_0(n, x)} = \tilde{M}_\alpha(n, x)$ .

6. Далее, можно явно выписать распределение сл.в.  $\tilde{\mathcal{T}}_0(X'_0, X''_0) \stackrel{\text{def}}{=} \max(\tilde{\mathcal{T}}_0(X'_0), \tilde{\mathcal{T}}_0(X''_0))$ :

$$\mathbf{P} \left\{ \tilde{\mathcal{T}}_0(X'_0, X''_0) \leq s \right\} = \begin{cases} 1 & \text{при } n' = n'' = 1; \\ 1 - e^{-\mu s} & \text{при } n' + n'' = 3; \\ (1 - e^{-\mu s})^2 & \text{при } n' = n'' = 2. \end{cases}$$

То есть

$$\mathbf{E} e^{\alpha \tilde{T}_0(X'_0, X''_0)} = \begin{cases} 1, & \text{при } n' = n'' = 1; \\ \frac{\mu}{\mu - \alpha}, & \text{при } n' + n'' = 3; \\ \frac{2\mu^2}{(2\mu - \alpha)(\mu - \alpha)}, & \text{при } n' = n'' = 2, \end{cases}$$

и

$$\begin{aligned} \int_{\mathcal{X}} \mathbf{E} e^{\alpha \tilde{T}_0(X'_0, X''_0)} \mathcal{P}(dX''_0) &\leq \frac{\mu}{\mu - \alpha} \left( k_\Gamma + \frac{2\mu(1 - k_\Gamma)}{2\mu - \alpha} \right) = \\ &= \frac{\mu^2(2\mu + 2\lambda - \alpha)}{(\mu - \alpha)(\lambda + \mu)(2\mu - \alpha)} \stackrel{\text{def}}{=} \mathcal{W}(\alpha, \lambda, \mu); \end{aligned}$$

эта оценка может быть уточнена.

7. В случае если по схеме, описанной в 6.3.5, после момента  $\mathcal{T}_k$  произошло склеивание процессов  $Z_t^i$  и  $Z_t''$  в момент  $\mathcal{T}_{k+1}$ , то разность  $\mathcal{T}_{k+1} - \mathcal{T}_k$  оценим сверху не величиной  $R^* + \hat{\xi}_k^{(1)}$ , а величиной  $\varrho \stackrel{\text{def}}{=} (\max(\xi_k^{(1)} + \xi_k^{(2)}, \xi_\ell^{(1)}) + \xi_{k'}^{(1)})$  – здесь мы учитываем, что  $\hat{\xi}_k^{(1)} \stackrel{\mathcal{D}}{=} \xi_k^{(1)}$ .

Ф.р.  $F_{1+2}(s) = \mathbf{P} \left\{ \xi_k^{(1)} + \xi_k^{(2)} \leq s \right\} = 1 - \frac{\lambda e^{-\mu s} - \mu e^{-\lambda s}}{\lambda - \mu}$ , и при  $\alpha < \min(\lambda, \mu)$

$$\mathbf{E} e^{\alpha \varrho} = \frac{2\lambda^5\mu - 2\lambda^3\mu^3 + 3\alpha(\lambda^3\mu^2 - \lambda^4\mu)}{(\lambda - \mu)(\lambda - \alpha)(2\lambda - \alpha)(\mu - \alpha)(\lambda + \mu - \alpha)} \stackrel{\text{def}}{=} \mathcal{A}(\alpha, \lambda, \mu).$$

8. Далее, следуя по пути доказательства теоремы 1, имеем: для сл.в.  $\nu$  такой, что  $\mathbf{P}\{\nu = k\} = pq^k$  ( $k \in \mathbb{Z}_+$ )

$$\begin{aligned} \mathbf{E} e^{\alpha T_{Z'_0, Z''_0}} &\leq \sum_{k=0}^{\infty} \mathbf{E} e^{\alpha(\tilde{T}_k + \varrho + \xi^{(1)})} \mathbf{P}\{\nu = k\} \leq \\ &\leq \sum_{k=0}^{\infty} \mathbf{E} e^{\alpha \tilde{T}_0} \left( \mathbf{E} e^{\alpha \xi^{(1)}} \right)^{k+1} \left( \mathbf{E} e^{\alpha \xi^{(2)}} \right)^k \mathbf{E} e^{\alpha \varrho} q^k = \\ &= \frac{\mathbf{E} e^{\alpha \tilde{T}_0} \mathcal{A}(\alpha, \lambda, \mu) \lambda}{(\lambda - \alpha) \left( 1 - \frac{q\lambda\mu}{(\lambda - \alpha)(\mu - \alpha)} \right)}; \end{aligned}$$

интегрируя это выражение по стационарной мере  $\mathcal{P}$ , получаем равномерную оценку:

$$\mathbb{E} e^{\alpha T_{Z'_0, Z''_0}} \leq \frac{\lambda \mu \mathcal{W}(\alpha, \lambda, \mu) \mathcal{A}(\alpha, \lambda, \mu)}{(\lambda - \alpha) \left( \mu - \alpha - \frac{q \lambda \mu}{(\lambda - \alpha)} \right)} \stackrel{\text{def}}{=} \mathfrak{K}(\alpha),$$

и для всех  $\alpha$ , удовлетворяющих неравенству

$$(10) \quad (\lambda - \alpha)(\mu - \alpha) > q \lambda \mu$$

верна оценка:  $|k_{\Gamma}(t) - k_{\Gamma}| < \mathfrak{K}(\alpha) e^{-\alpha t}$ .

9. Пусть  $\lambda = 1, \mu = 5$ . Тогда  $\mathcal{W}(\alpha, 1, 5) = \frac{25(12 - \alpha)}{6(10 - \alpha)(5 - \alpha)}$

$q = \frac{5}{12}, p = \frac{7}{12}$ ; условие (10) выполнено при  $\alpha < 0,370044$ .

При  $\alpha = 0,3$ :  $\mathcal{A}(\alpha, \lambda, \mu) \approx 1,7408979$ ;  $\mathcal{W}(\alpha, \lambda, \mu) \approx 1,06931345$ ;  $\mathfrak{K}(0,3) \approx 10,3882007$ , т.е. при любых начальных условиях  $|k_{\Gamma}(t) - k_{\Gamma}| < 10,3882007 e^{-0,3t}$ , что хуже (1), но даёт вполне приемлемую оценку скорости сходимости, существенно лучшую, чем в примерах 1 и 2.

#### 4.4. ВЫВОДЫ

Несмотря на то, что оценка (8), как будет видно из доказательства теоремы 1, строится довольно грубо, полученные формулы позволяют давать оценки скорости сходимости коэффициента готовности для достаточно широкого класса распределений. Однако диапазон изменения  $\alpha$ , при котором возможно применение оценки (8), как правило, довольно мал.

Численные эксперименты показали, что чем выше скорость сходимости  $k_{\Gamma}(t)$  к предельному значению, тем хуже полученные по формуле (8) оценки.

Однако, как видно из примера 3, знание некоторой дополнительной информации о функциях распределения  $F_1(s)$  и  $F_2(s)$  позволяет улучшить оценку скорости сходимости, избежав излишних округлений неравенств в оценках доказательства теоремы 1.

При этом из таблиц 2 и 4 видно, что при наблюдении работы реального восстанавливаемого элемента для получения приемлемой оценки  $k_{\Gamma}$  достаточно наблюдать около 100 периодов регенерации.

Таблицы 2–4 показывают, что, как правило, вычисленная по формуле (8) оценка для  $K(\alpha)$  отличается от эмпирического значения в  $10^2$ – $10^3$  раз.

Ниже приводится доказательство основной теоремы 1.

## **5. Предварительные соображения и вспомогательные факты**

### **5.1. ОБ ИССЛЕДУЕМОМ ПРОЦЕССЕ**

Мы рассматриваем процесс  $X_t$ , который является регенерирующим процессом. В дальнейшем в качестве периодов регенерации мы будем рассматривать интервалы вида  $(t_k, t_{k+1})$  – интервалы между последовательными попаданиями процесса  $X_t$  в состояние  $(1, 0)$  (вообще говоря, можно рассматривать и другие точки регенерации). Моменты  $(t_1, t_2, \dots)$  (или  $(t'_1, t'_2, \dots)$ ) – точки его регенерации, а длина периода регенерации  $\xi \stackrel{\text{def}}{=} \xi^{(1)} + \xi^{(2)}$  имеет математическое ожидание, равное  $\mathbf{E} \xi^{(1)} + \mathbf{E} \xi^{(2)} = T_1 + T_2$ .

Мы будем использовать также сл.в.

$$(11) \quad \zeta_t^* \stackrel{\text{def}}{=} (\inf\{T > t : X_t = (1, 0)\} - t),$$

$\zeta_t^*$  – время до начала следующего периода регенерации («время перескока» периода регенерации) процесса  $X_t$ .

Известно, что процесс  $X_t$  является строго марковским (см., например, [6]).

### **5.2. Полезные следствия из Ключевой теоремы восстановления (теоремы Смита)**

Следствием из Ключевой теоремы восстановления ([5, С.23]) являются следующие факты:

**Предложение 1.** *В условиях теоремы 1 распределение процесса  $X_t$  при любых начальных условиях слабо сходится к распределению  $\mathcal{P}$ , которое задаётся следующим образом (см. [4,*

§ 2.6], [9] и формулу (6):

$$\mathcal{P}(\{n = 1, x \leq s\}) = k_{\Gamma} \int_0^s \tilde{f}^{(1)}(u) du;$$

$$\mathcal{P}(\{n = 2, x \leq s\}) = (1 - k_{\Gamma}) \int_0^s \tilde{f}^{(2)}(u) du,$$

где  $k_{\Gamma} = \frac{T_1}{T_1 + T_2}$ ,  $\tilde{f}^{(1)}(u)$  определены в (6).  $\triangleright$

**Предложение 2.** Для процесса восстановления  $t_0 = 0$ ,  $t_1, t_2, \dots, t_k \dots, t_{k+1} - t_k = \xi$  - н.о.р. сл.в. с нерешётчатой ф.р.  $F(x)$ , имеющие конечный второй момент, верно следующее неравенство:

$$\mathbf{E} x_t^* \leq \lim_{t \rightarrow \infty} \mathbf{E} x_t^* = \frac{\mathbf{E} \xi^2}{2\mathbf{E} \xi},$$

где  $x_t^* \stackrel{\text{def}}{=} (\min\{t_k : t_k > t\} - t)$  - время перескока ([2, 3]) процесса восстановления (или прямое время восстановления - [5]) в момент  $t$  (см., например, [9]).  $\triangleright$

**Следствие 2.** В условиях теоремы 1

$$\mathbf{E}(\zeta_t^* | t > t_1) \leq \lim_{t \rightarrow \infty} \mathbf{E} \zeta_t^* = \Theta_1 \stackrel{\text{def}}{=} \frac{\mathbf{E}(\xi)^2}{2\mathbf{E} \xi} = \frac{\mathbf{E}(\xi^{(1)} + \xi^{(2)})^2}{2(T_1 + T_2)}. \triangleright$$

**Следствие 3.** В условиях теоремы 1

$$(12) \quad \mathbf{E}(\zeta_t^* | t > t_1, X_t \in \mathcal{S}_2) \leq \Theta \stackrel{\text{def}}{=} \frac{\mathbf{E}(\xi^{(2)})^2}{2T_2},$$

где  $\zeta_t^* \stackrel{\text{def}}{=} (\inf\{s > t : Z_s'' = (1, 0)\} - t)$  - время от момента  $t$  до попадания процесса  $X_t$  в множество  $\mathcal{S}_1$ .  $\triangleright$

### 5.3. ОБОЗНАЧЕНИЯ

5.3.1.  $\mathcal{L}$  - мера Лебега.

5.3.2.  $\mathcal{U}_{\Omega}$  - равномерно распределённая случайная величина, заданная на  $\Omega = [0; 1]$ .

$$5.3.3. \varphi_{x,y}(s) \stackrel{\text{def}}{=} \min \left( f_x^{(1)}(s), f_y^{(1)}(s) \right); \varkappa_{x,y} \stackrel{\text{def}}{=} \int_0^{\infty} \varphi_{x,y}(s) ds.$$

**Замечание 3.** В условиях теоремы 1

$$\varphi_{x,y}(s) \geqslant ce^{-Cs} \quad \text{и} \quad \varkappa_{x,y} \geqslant \frac{c}{C}. \triangleright$$

5.3.4. Для монотонно неубывающей функции  $\Psi(s)$  обозначим

$$\Psi^{-1}(u) \stackrel{\text{def}}{=} \sup\{v : \Psi(v) \leqslant u\}.$$

**Замечание 4.** Если  $\Psi(s)$  – функция распределения сл.в., то  $\Psi^{-1}(\mathcal{U}_\Omega)$  – сл.в. на  $\Omega$  с ф.р.  $\Psi(s)$ .  $\triangleright$

$$5.3.5. \Phi_{x,y}(s) \stackrel{\text{def}}{=} \int_0^s \varphi_{x,y}(u) du, \quad \widehat{\Phi}_{x,y}(s) \stackrel{\text{def}}{=} F_x^{(1)}(s) - \Phi_{x,y}(s).$$

**Замечание 5.**  $\Phi_{x,y}(s)$  и  $\widehat{\Phi}_{x,y}(s)$  – неотрицательные неубывающие непрерывные функции и

$$\Phi_{x,y}(+\infty) = \varkappa_{x,y}, \quad \widehat{\Phi}_{x,y}(+\infty) = 1 - \varkappa_{x,y},$$

$$\Phi_{x,y}(0+) = \widehat{\Phi}_{x,y}(0+) = 0. \triangleright$$

$$5.3.6. \Xi_{x,y}(s) \stackrel{\text{def}}{=} \begin{cases} \Phi_{x,y}^{-1}(s) & \text{при } s \in [0; \varkappa_{x,y}); \\ \widehat{\Phi}_{x,y}^{-1}(s - \varkappa_{x,y}) & \text{при } s \in [\varkappa_{x,y}; 1). \end{cases}$$

**Замечание 6.**  $\Xi_{x,y}(\mathcal{U}_\Omega)$  – сл.в. на  $\Omega$  с ф.р.  $F_x^{(1)}(s)$ . При этом  $\mathbf{P}\{\Xi_{x,y}(\mathcal{U}_\Omega) = \Xi_{y,x}(\mathcal{U}_\Omega)\} = \varkappa_{x,y}$  – подробнее об этом см. [8, 11, 12].  $\triangleright$

Кроме того, в условиях теоремы 1 плотность  $\widetilde{p}_{x,y}(s)$  распределения сл.в.  $(\Xi_{x,y}(\mathcal{U}_\Omega) | \Xi_{x,y}(\mathcal{U}_\Omega) = \Xi_{y,x}(\mathcal{U}_\Omega))$  удовлетворяет неравенству  $\widetilde{p}_{x,y}(s) \leqslant Ce^{-cs}$  (см. условие (7)).

## 6. Доказательство теоремы 1

### 6.1. КОНСТРУКЦИЯ ПРОЦЕССА $X_t$

Пусть  $(\Omega, \mathcal{F}, \mathbf{P}) = ([0, 1], \mathcal{B}([0, 1]), \mathcal{L})$ . Возьмём прямое произведение счётного множества экземпляров пространства



$(\Omega, \mathcal{F}, \mathbf{P})$ ), предварительно их пронумеровав:

$$(\Omega^X, \mathcal{F}^X, \mathbf{P}^X) = (\Omega_0, \mathcal{F}_0, \mathbf{P}_0) \times (\Omega_1, \mathcal{F}_1, \mathbf{P}_1) \times (\Omega_2, \mathcal{F}_2, \mathbf{P}_2) \times \dots \\ \dots \times (\Omega_k, \mathcal{F}_k, \mathbf{P}_k) \times (\Omega_{k+1}, \mathcal{F}_{k+1}, \mathbf{P}_{k+1}) \times \dots$$

и будем строить на нём процесс  $X_t$ , выходящий из состояния  $X_0 = (i, a)$ , который мы обозначим  $X_t^{(i,a)}$ , следующим образом.

Используя распределения (3) и (4), положим:

$\theta_0 \stackrel{\text{def}}{=} 0$ ,  $i_0 \stackrel{\text{def}}{=} i$ ;  $\theta_1 \stackrel{\text{def}}{=} \xi^{(i_0)}(a) + \theta_0$ , где  $\xi^{(i_0)}(a)$  – случайная величина с функцией распределения  $F_a^{(i)}(s)$  на пространстве  $(\Omega_0, \mathcal{F}_0, \mathbf{P}_0)$ ;

$i_1 \stackrel{\text{def}}{=} i_0 - (-1)^{i_0}$ ,  $\theta_2 \stackrel{\text{def}}{=} \xi_1^{(i_1)} + \theta_1$ , где  $\xi_1^{(i_1)}$  – случайная величина с функцией распределения  $F^{(i_1)}(s)$  на пространстве  $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1)$ ;

$i_2 \stackrel{\text{def}}{=} i_1 - (-1)^{i_1}$ ;  $\theta_3 \stackrel{\text{def}}{=} \xi_2^{(i_2)} + \theta_1$ , где  $\xi_2^{(i_2)}$  – случайная величина с функцией распределения  $F^{(i_2)}(s)$  на пространстве  $(\Omega_2, \mathcal{F}_2, \mathbf{P}_2)$ ;

· · · и т.д.; · · ·

$i_k \stackrel{\text{def}}{=} i_{k-1} - (-1)^{i_{k-1}}$ ;  $\theta_{k+1} \stackrel{\text{def}}{=} \xi_k^{(i_k)} + \theta_1$ , где  $\xi_k^{(i_k)}$  – случайная величина с функцией распределения  $F^{(i_k)}(s)$  на пространстве  $(\Omega_k, \mathcal{F}_k, \mathbf{P}_k)$ ;

· · · и т.д. · · ·

Обозначим  $r(t) \stackrel{\text{def}}{=} \max\{k : \theta_k < t\}$  и  $U_t \stackrel{\text{def}}{=} (i_{r(t)}, t - \theta_{r(t)})$ .

Построенный процесс  $U_t$  начинается из состояния  $(i, a)$ , первое изменение компоненты  $n_t$  этого процесса происходит в момент, когда заканчивается остаточное время пребывания элемента в состоянии  $i$  при условии, что к начальному моменту времени элемент находился в этом состоянии  $a$  единиц времени; в дальнейшем элемент поочерёдно находится в состояниях  $j = i - (-1)^i$  и  $i$  в течение случайного времени с функциями распределения  $F^{(j)}(s)$  и  $F^{(i)}(s)$  соответственно;  $\theta_k$  – моменты скачков (изменений) первой компоненты процесса  $U_t$ .

То есть  $(U_t, t \geq 0) \stackrel{\mathcal{D}}{=} (X_t^{(i,a)}, t \geq 0)$ , где  $X_t^{(i,a)}$  – процесс  $X_t$  с начальным состоянием  $(i, a)$ .

## 6.2. КОНСТРУКЦИЯ СКЛЕЙКИ $(Z'_t, Z''_t)$

Для конструирования *одного* процесса  $X_t$  на пространстве  $(\Omega^X, \mathcal{F}^X, \mathbf{P}^X)$  строятся случайные моменты  $\theta_0, \theta_1, \dots$ , в которые происходит изменение состояния (исправное или неисправное) восстанавливаемого элемента. Для конструирования пары процессов нам нужны две последовательности таких случайных моментов. В предложенной ниже конструкции будут строиться случайные моменты  $\vartheta_0 = 0, \vartheta_1, \vartheta_2, \dots$ , в которые происходит изменение состояния *хотя бы одного* из пары процессов (при этом конструкция будет такова, что в некоторые моменты может произойти изменение состояния *обоих* процессов).

Пару процессов  $(Z'_t, Z''_t) = ((m'_t, z'_t), (m''_t, z''_t))$  будем строить на вероятностном пространстве

$$\begin{aligned} (\Omega^{Z', Z''}, \mathcal{F}^{Z', Z''}, \mathbf{P}^{Z', Z''}) &\stackrel{\text{def}}{=} \prod_{k=0}^{\infty} ((\Omega'_k, \mathcal{F}'_k, \mathbf{P}'_k) \times (\Omega''_k, \mathcal{F}''_k, \mathbf{P}''_k)) = \\ &= \prod_{k=0}^{\infty} (\Omega'_k, \mathcal{F}'_k, \mathbf{P}'_k) \times \prod_{k=0}^{\infty} (\Omega''_k, \mathcal{F}''_k, \mathbf{P}''_k) = \\ &= (\Omega^{Z'}, \mathcal{F}^{Z'}, \mathbf{P}^{Z'}) \times (\Omega^{Z''}, \mathcal{F}^{Z''}, \mathbf{P}^{Z''}), \end{aligned}$$

здесь  $(\Omega'_k, \mathcal{F}'_k, \mathbf{P}'_k)$  и  $(\Omega''_k, \mathcal{F}''_k, \mathbf{P}''_k)$  – экземпляры пространства  $(\Omega, \mathcal{F}, \mathbf{P})$ .

Пусть

$$\begin{aligned} Z'_0 = (m'_0, z'_0) &= (n'_0, x'_0) = X'_0, \text{ и} \\ Z''_0 = (m''_0, z''_0) &= (n''_0, x''_0) = X''_0. \end{aligned}$$

Положим  $\vartheta_0 \stackrel{\text{def}}{=} 0$ . Конструирование моментов  $\vartheta_k$  строится последовательно, в некотором смысле «по индукции».

Пусть момент  $\vartheta_k$  уже определён с помощью сл.в., заданных на пространстве  $\prod_{\ell=0}^{k-1} (\Omega'_\ell, \mathcal{F}'_\ell, \mathbf{P}'_\ell) \times \prod_{\ell=0}^{k-1} (\Omega''_\ell, \mathcal{F}''_\ell, \mathbf{P}''_\ell)$ , и в этот момент  $Z'_{\vartheta_k} = (m'_{\vartheta_k}, z'_{\vartheta_k})$  и  $Z''_{\vartheta_k} = (m''_{\vartheta_k}, z''_{\vartheta_k})$ .

6.2.1. Если  $Z'_{\vartheta_k} = Z''_{\vartheta_k}$ , то полагаем  $Z'_t = Z'_t$  при  $t \geq \vartheta_k$  и строим процесс  $Z'_t$  для  $t > \vartheta_k$  на пространствах  $(\Omega'_k, \mathcal{F}'_k, \mathbf{P}'_k)$ ,  $(\Omega'_{k+1}, \mathcal{F}'_{k+1}, \mathbf{P}'_{k+1})$ , ... по схеме, описанной в разделе 6.1.

6.2.2. Если  $Z'_{\vartheta_k} \neq Z''_{\vartheta_k}$  и  $m'_{\vartheta_k} + m''_{\vartheta_k} > 2$ , т.е. не выполнено  $\left\{ Z'_{\vartheta_k} \in \mathcal{S}_1, Z''_{\vartheta_k} \in \mathcal{S}_1 \right\}$ , полагаем  $\vartheta_{k+1} \stackrel{\text{def}}{=} \vartheta_k + \min(\vartheta'_{k+1}, \vartheta''_{k+1})$ , где

$$\vartheta'_{k+1} \stackrel{\text{def}}{=} \left( F_{z'_{\vartheta_k}}^{(m'_{\vartheta_k})} \right)^{-1} \left( \mathcal{U}_{\Omega'_k} \right), \quad \vartheta''_{k+1} \stackrel{\text{def}}{=} \left( F_{z''_{\vartheta_k}}^{(m''_{\vartheta_k})} \right)^{-1} \left( \mathcal{U}_{\Omega''_k} \right).$$

Для  $t \in [\vartheta_k; \vartheta_{k+1})$

(13)

$$Z'_t \stackrel{\text{def}}{=} (m'_{\vartheta_k}, z'_{\vartheta_k} + (t - \vartheta_k)), \quad Z''_t \stackrel{\text{def}}{=} (m''_{\vartheta_k}, z''_{\vartheta_k} + (t - \vartheta_k));$$

$$Z'_{\vartheta_{k+1}} \stackrel{\text{def}}{=} \begin{cases} (m'_{\vartheta_k}, z'_{\vartheta_k} + (t - \vartheta_{k+1})) & \text{при } \vartheta'_{k+1} \neq \min(\vartheta'_{k+1}, \vartheta''_{k+1}); \\ (m'_{\vartheta_k} - (-1)^{m'_{\vartheta_k}}, 0) & \text{при } \vartheta'_{k+1} = \min(\vartheta'_{k+1}, \vartheta''_{k+1}); \end{cases}$$

$$Z''_{\vartheta_{k+1}} \stackrel{\text{def}}{=} \begin{cases} (m''_{\vartheta_k}, z''_{\vartheta_k} + (t - \vartheta_{k+1})) & \text{при } \vartheta''_{k+1} \neq \min(\vartheta'_{k+1}, \vartheta''_{k+1}); \\ (m''_{\vartheta_k} - (-1)^{m''_{\vartheta_k}}, 0) & \text{при } \vartheta''_{k+1} = \min(\vartheta'_{k+1}, \vartheta''_{k+1}). \end{cases}$$

Здесь  $\vartheta'_{k+1}$  и  $\vartheta''_{k+1}$  – остаточные времена пребывания процессов  $Z'_t$  и  $Z''_t$  в состояниях  $m'_{\vartheta_k}$  и  $m''_{\vartheta_k}$  в момент времени  $\vartheta_k$  соответственно; в момент  $\vartheta_{k+1}$  один из них изменяет это состояние; может случиться так, что состояние в момент  $\vartheta_{k+1}$  изменят оба процесса.

6.2.3. Если  $m'_{\vartheta_k} = m''_{\vartheta_k} = 1$  и  $Z'_{\vartheta_k} \neq Z''_{\vartheta_k}$ , полагаем:  $\vartheta'_{k+1} \stackrel{\text{def}}{=} \Xi_{z'_{\vartheta_k}, z''_{\vartheta_k}}(\mathcal{U}_{\Omega'_k})$ ,  $\vartheta''_{k+1} \stackrel{\text{def}}{=} \Xi_{z''_{\vartheta_k}, z'_{\vartheta_k}}(\mathcal{U}_{\Omega'_k})$ ,  $\vartheta_{k+1} \stackrel{\text{def}}{=} \vartheta_k + \min(\vartheta'_{k+1}, \vartheta''_{k+1})$ . Здесь опять  $\vartheta'_{k+1}$  и  $\vartheta''_{k+1}$  – остаточные времена пребывания процессов  $Z'_t$  и  $Z''_t$  в состояниях  $m'_{\vartheta_k}$  и  $m''_{\vartheta_k}$  соот-

ответственно; в момент  $\vartheta_{k+1}$  один из них изменяет состояние. При этом  $\mathbf{P}\{\vartheta'_{k+1} = \vartheta''_{k+1}\} = \kappa_{z'_{\vartheta_k}, z''_{\vartheta_k}} \geq \frac{c}{C}$ .

То есть с вероятностью  $\kappa_{z'_{\vartheta_k}, z''_{\vartheta_k}} \geq \frac{c}{C}$  процессы  $Z'_t$  и  $Z''_t$  одновременно перейдут в множество  $\mathcal{S}_2$  (в соответствии с Замечанием 6), тогда  $Z'_{\vartheta_{k+1}} = Z''_{\vartheta_{k+1}} = (2, 0)$  и, в соответствии с п. 1 этой конструкции, при  $t > \vartheta_{k+1}$   $Z'_t = Z''_t$ . Поведение процессов  $Z'_t$  и  $Z''_t$  на промежутке  $[\vartheta_k, \vartheta_{k+1}]$  задаётся с помощью формул (13).

**Замечание 7.** Обозначим  $\widehat{\xi}_{k+1} \stackrel{\text{def}}{=} \left\{ \vartheta'_{k+1} \Big|_{\vartheta'_{k+1} = \vartheta''_{k+1}} \right\}$ . Сл.в.  $\widehat{\xi}_{k+1}$  имеет плотность  $\varphi_{z'_{\vartheta_k}, z''_{\vartheta_k}}(s) \in [ce^{-Cs}; Ce^{-cs}]$ , т.е.  $\mathbf{E} e^{\alpha \widehat{\xi}_{k+1}} \leq \frac{C}{c - \alpha}$  при  $\alpha < c$ .  $\triangleright$

**Замечание 8.** Можно показать, что построенная таким образом пара процессов  $(Z'_t, Z''_t)$  является марковским процессом, и распределение компонент этого процесса совпадает с распределением процессов  $X'_t = X_t^{X'_0}$  и  $X''_t = X_t^{X''_0}$ :  $(X_t^{X'_0}, t \geq 0) \stackrel{\mathcal{D}}{=} (Z'_t, t \geq 0)$  и  $(X_t^{X''_0}, t \geq 0) \stackrel{\mathcal{D}}{=} (Z''_t, t \geq 0)$ . В настоящей статье, в связи с громоздкостью доказательства этого факта, мы его не приводим.  $\triangleright$

Теперь наша задача – оценить распределение (и экспоненциальные моменты) сл.в.  $T_{Z'_0, Z''_0}$ , т.е. убедиться в успешности склейки  $(Z'_t, Z''_t)$ .

### 6.3. ОЦЕНКА РАСПРЕДЕЛЕНИЯ $T_{Z'_0, Z''_0}$

6.3.1. Пусть

$$(14) \quad \begin{aligned} \mathcal{T}'_0 &\stackrel{\text{def}}{=} \inf\{t > 0 : Z'_t = (1, 0)\}, \\ \mathcal{T}''_0 &\stackrel{\text{def}}{=} \inf\{t > 0 : Z''_t = (1, 0)\}; \end{aligned}$$

положим  $\mathcal{T}_0 \stackrel{\text{def}}{=} \max(\mathcal{T}'_0, \mathcal{T}''_0) \leq (\mathcal{T}'_0 + \mathcal{T}''_0) \stackrel{\text{def}}{=} \widehat{T}_0$ ;  $M \stackrel{\text{def}}{=} \mathbf{E} \widehat{T}_0$ .

Здесь  $\mathcal{T}'_0$  и  $\mathcal{T}''_0$  – моменты начала *полных* периодов регенерации процессов  $Z'_t$  и  $Z''_t$ , поэтому после момента  $\mathcal{T}_0$  поведение обоих процессов не зависит от начального состояния. До момента  $\mathcal{T}_0$  процессы могли склеиться (например, в случае, ес-

ли  $X'_0, X''_0 \in \mathcal{S}_1$ ), но мы отбросим эту возможность, что сделает нашу оценку распределения  $\Gamma_{Z'_0, Z''_0}$  более грубой.

Не ограничивая общности, будем считать, что  $Z'_0 = (1, 0)$  (случай  $Z''_0 = (1, 0)$  рассматривается аналогично).

Обозначим  $\mathcal{T}_k \stackrel{\text{def}}{=} \inf\{t > \mathcal{T}_{k-1} : Z'_t = (1, 0)\}$  – моменты начала периодов регенерации процесса  $Z'_t$  начиная с момента  $\mathcal{T}_0$  ( $k \in \mathbb{N}$ );  $\mathcal{T}_{k+1} - \mathcal{T}_k = \xi_k^{(1)} + \xi_k^{(2)}$  ( $k \in \mathbb{Z}_+$ ).

6.3.2. В момент времени  $\mathcal{T}_k$  начинается период работы элемента длиной  $\xi_k^{(1)}$ . В этот момент (если ранее не произошло совпадения (склейки) процессов  $Z'_t$  и  $Z''_t$ ) процесс  $Z''_t$  находится на некотором периоде регенерации, и он находится в исправном состоянии ( $Z''_{\mathcal{T}_k} \in \mathcal{S}_1$ ) с вероятностью  $k_{\Gamma}^{X''}(\mathcal{T}_k)$ , или в неисправном ( $Z''_{\mathcal{T}_k} \in \mathcal{S}_2$ ) состоянии с вероятностью  $(1 - k_{\Gamma}^{X''}(\mathcal{T}_k))$ . Рассмотрим оба случая.

6.3.3.  $Z''_{\mathcal{T}_k} \in \mathcal{S}_1$ . При  $n(Z''_{\mathcal{T}_k}) = 1$ , в соответствии с конструкцией раздела 6.2, совместное пребывание в множестве  $\mathcal{S}_1$  с вероятностью  $\varkappa_{0, z''_{\mathcal{T}_k}} \geq \frac{c}{C}$  может завершиться склейкой процессов  $Z'_t$  и  $Z''_t$  через время  $\hat{\xi}_k \stackrel{\text{def}}{=} \left( \Xi_{0, z''_{\mathcal{T}_k}}(\mathcal{U}_{\Omega}) | \Xi_{0, z''_{\mathcal{T}_k}}(\mathcal{U}_{\Omega}) = \Xi_{Z''_{\mathcal{T}_k}, 0}(\mathcal{U}_{\Omega}) \right)$ .

6.3.4.  $Z''_{\mathcal{T}_k} \in \mathcal{S}_2$ . Если  $Z''_{\mathcal{T}_k} \in \mathcal{S}_2$ , рассмотрим сл.в.  $\varsigma_{\mathcal{T}_k}^*$  – остаточное время пребывания  $Z''_t$  в множестве  $\mathcal{S}_2$ ;  $\varsigma_{\mathcal{T}_k}^* \stackrel{\text{def}}{=} (\inf\{t > \mathcal{T}_k : Z''_t = (1, 0)\} - \mathcal{T}_k)$  и оценим  $\mathbf{P}\{\varsigma_{\mathcal{T}_k}^* < \xi^{(1)}\}$ ; событие  $\mathcal{E}_k \stackrel{\text{def}}{=} \{\varsigma_{\mathcal{T}_k}^* < \xi^{(1)}\}$  означает, что на интервале  $(\mathcal{T}_k, \mathcal{T}_k + \xi_k^{(1)})$  процессы  $Z'_t$  и  $Z''_t$  одновременно окажутся в множестве  $\mathcal{S}_1$  (после чего они могут склеиться с ненулевой вероятностью).

Поскольку  $\mathbf{E}\varsigma_{\mathcal{T}_k}^* \leq \Theta$  (12), оценим вероятность события  $\mathcal{E}_k^R \stackrel{\text{def}}{=} \{\varsigma_{\mathcal{T}_k}^* < R < \xi^{(1)}\}$ , где  $R$  – некоторое число. Очевидно,  $\mathcal{E}_k^R \subset \mathcal{E}_k$ .

Пусть  $R > \Theta$ ;  $\mathbf{P}\{\xi_k^{(1)} > R\} = 1 - F^{(1)}(R) \geq e^{-CR}$ , а по неравенству Маркова  $\mathbf{P}\{\varsigma_{\mathcal{T}_k}^* < R\} = 1 - \mathbf{P}\{\varsigma_{\mathcal{T}_k}^* \geq R\} \leq 1 - \frac{\Theta}{R}$ .

Поэтому

$$\mathbf{P}(\mathcal{E}_k) \geq \mathbf{P}(\mathcal{E}_k^R) \geq \rho = \rho(C, R, \Theta) \stackrel{\text{def}}{=} e^{-CR} \left(1 - \frac{\Theta}{R}\right).$$

Вероятность  $\rho$  достигает максимума при  $R = R^*(C, \Theta)$ , т.е.  $\mathbf{P}(\mathcal{E}_k) \geq \rho$ .

Итак, на интервале  $(\mathcal{T}_k, \mathcal{T}_k + R^*)$  оба процесса окажутся в множестве  $\mathcal{S}_1$  с вероятностью большей, чем  $\rho$ .

6.3.5.  $\{Z''_{\mathcal{T}_k} \in \mathcal{S}_2\}, \mathcal{E}_k$ . Если произошло событие  $\mathcal{E}_k$ , и в момент  $\widehat{\mathcal{T}}_k \stackrel{\text{def}}{=} \mathcal{T}_k + \varsigma_{\mathcal{T}_k}^*$  выполнено  $Z'_{\widehat{\mathcal{T}}_k} \in \mathcal{S}_1, Z''_{\widehat{\mathcal{T}}_k} \in \mathcal{S}_1$ , то, в соответствии с конструкцией раздела 6.2, с вероятностью  $\varkappa_{0, \varsigma_{\mathcal{T}_k}^*} \geq \frac{c}{C}$  в момент  $\widehat{\mathcal{T}}_k \stackrel{\text{def}}{=} \mathcal{T}_k + \xi_k^{(1)}$  процессы  $Z'_t$  и  $Z''_t$  совпадут, при этом (см. замечание 7)

$$\xi_k^{(1)} \left| \left\{ \mathcal{E}_k, \left\{ Z'_{\widehat{\mathcal{T}}_k} = Z''_{\widehat{\mathcal{T}}_k} \right\} \right\} \right. = \varsigma_{\mathcal{T}_k}^* + \widehat{\xi}_k \leq R^* + \widehat{\xi}_k.$$

6.3.6. Поэтому с вероятностью большей, чем

$$(15) \quad \frac{c}{C} \left( k_{\Gamma}^{X''_0}(\mathcal{T}_k) + \left(1 - k_{\Gamma}^{X''_0}(\mathcal{T}_k)\right) e^{-CR} \left(1 - \frac{\Theta}{R}\right) \right) > \\ > \frac{c e^{-CR^*(C, \Theta)}}{C} \left(1 - \frac{\Theta}{R^*}\right) \stackrel{\text{def}}{=} 1 - q,$$

процессы  $Z'_t$  и  $Z''_t$  совпадут в момент времени

$$(16) \quad T_{Z'_0, Z''_0} \leq \mathcal{T}_k + R^* + \widehat{\xi}_k.$$

6.3.7. Итак, учитывая, что начиная с момента  $T_{Z'_0, Z''_0}$  распределения процессов  $Z'_t$  и  $Z''_t$  совпадают, для всех  $A \in \mathcal{B}(\mathcal{X})$  верно основное неравенство склеивания (basic coupling inequality):

$$\left| \mathcal{P}_t^{X'_0}(A) - \mathcal{P}_t^{X''_0}(A) \right| = \\ = \left| \mathbf{P}\{X'_t \in A\} - \mathbf{P}\{X''_t \in A\} \right| = \left| \mathbf{P}\{Z'_t \in A\} - \mathbf{P}\{Z''_t \in A\} \right| = \\ = \left| \mathbf{P}\{Z'_t \in A, T_{Z'_0, Z''_0} \leq t\} + \mathbf{P}\{Z'_t \in A, T_{Z'_0, Z''_0} > t\} - \right. \\ \left. - \mathbf{P}\{Z''_t \in A, T_{Z'_0, Z''_0} \leq t\} - \mathbf{P}\{Z''_t \in A, T_{Z'_0, Z''_0} > t\} \right| \leq$$

$$\begin{aligned} &\leq \left| \mathbf{P}\{Z'_t \in A, T_{Z'_0, Z''_0} > t\} - \mathbf{P}\{Z''_t \in A, T_{Z'_0, Z''_0} > t\} \right| \leq \\ &\leq \mathbf{P}\{T_{Z'_0, Z''_0} > t\} = \mathbf{P}\{e^{\alpha T_{Z'_0, Z''_0}} > e^{\alpha t}\} \leq \\ &\leq \frac{\mathbf{E} e^{\alpha T_{Z'_0, Z''_0}}}{e^{\alpha t}} \leq K(\alpha, Z'_0, Z''_0) e^{-\alpha t} = K(\alpha, X'_0, X''_0) e^{-\alpha t}, \end{aligned}$$

т.е.

$$\begin{aligned} \left\| \mathcal{P}_t^{X'_0}(A) - \mathcal{P}_t^{X''_0}(A) \right\|_{TV} &\stackrel{\text{def}}{=} \sup_{A \in \mathcal{B}(\mathcal{X})} \left| \mathcal{P}_t^{X'_0}(A) - \mathcal{P}_t^{X''_0}(A) \right| \leq \\ &\leq K(\alpha, X'_0, X''_0) e^{-\alpha t}. \end{aligned}$$

6.3.8. Оценим  $\mathbf{E} e^{\alpha T_{Z'_0, Z''_0}}$  при  $\alpha$ , удовлетворяющих (7).

Обозначим  $\nu$  – номер периода регенерации  $(\mathcal{T}_k; \mathcal{T}_{k+1}]$ , на котором произошло совпадение процессов  $Z'_t$  и  $Z''_t$ . Заметим, что при конструировании склейки  $(Z'_t, Z''_t)$  были схеме учтены не все возможные ситуации склеивания процессов. Учитывая (15) и (16),

$$\mathbf{P}\{\nu = k\} = \mathbf{P}\{\bar{\mathcal{E}}_0 \cap \bar{\mathcal{E}}_1 \cap \bar{\mathcal{E}}_2 \cap \dots \cap \bar{\mathcal{E}}_{k-1} \cap \mathcal{E}_k\} \leq q^k;$$

$$\begin{aligned} \mathbf{E} e^{\alpha T_{Z'_0, Z''_0}} &\leq \sum_{k=0}^{\infty} \mathbf{E} e^{\alpha(\mathcal{T}_k + R^* + \widehat{\xi}_k^{(1)})} \mathbf{P}\{\nu = k\} \leq \\ &\leq \sum_{k=0}^{\infty} \mathbf{E} e^{\alpha\left(\mathcal{T}_0 + \sum_{r=1}^k (\xi_r^{(1)} + \xi_r^{(2)}) + R^* + \widehat{\xi}_k^{(1)}\right)} q^k = \\ &= e^{\alpha R^*} \sum_{k=0}^{\infty} \mathbf{E} e^{\alpha \mathcal{T}_0} \mathbf{E} e^{\alpha \widehat{\xi}_k^{(1)}} \left( \mathbf{E} e^{\alpha \xi^{(1)}} \right)^k \left( \mathbf{E} e^{\alpha \xi^{(2)}} \right)^k q^k \leq \\ &\leq e^{\alpha R^*} e^{\alpha \mathbf{E} e^{\alpha \mathcal{T}_0}} \frac{C}{c - \alpha} \left( 1 - \mathbf{E} e^{\alpha \xi^{(1)}} \mathbf{E} e^{\alpha \xi^{(2)}} q \right)^{-1}. \end{aligned}$$

Далее,

$$\begin{aligned} \mathbf{E} e^{\alpha \mathcal{T}_0} &= \mathbf{E} e^{\alpha \mathcal{T}_0(Z'_0, Z''_0)} \leq \mathbf{E} e^{\alpha(\mathcal{T}'_0(Z'_0) + \mathcal{T}''_0(Z''_0))} = \\ &= \mathbf{E} e^{\alpha \mathcal{T}'_0(Z'_0)} \times \mathbf{E} e^{\alpha \mathcal{T}''_0(Z''_0)}; \quad \text{и} \quad \mathbf{E} e^{\alpha \mathcal{T}'_0(Z'_0)} = M_\alpha(m'_0, z'_0); \\ &\quad \mathbf{E} e^{\alpha \mathcal{T}''_0(Z''_0)} = M_\alpha(m''_0, z''_0) - \text{см. п. 2.1.7.} \end{aligned}$$

Поэтому (учитывая (5))

$$\mathbf{E} e^{\alpha T_{z'_0, z''_0}} \leq \frac{C e^{\alpha R^*} M_\alpha(Z'_0) M_\alpha(Z''_0)}{(c - \alpha) (1 - \mathbf{E} e^{\alpha \xi^{(1)}} \mathbf{E} e^{\alpha \xi^{(2)}} q)} \stackrel{\text{def}}{=} \\ \stackrel{\text{def}}{=} \mathfrak{K}(\alpha, c, C, Z'_0, Z''_0) = \mathfrak{K}(\alpha, Z'_0, Z''_0);$$

$$\int_{\mathcal{X}} M_\alpha(Z''_0) \mathcal{P}(dZ''_0) = \\ = \mathbf{E} e^{\alpha \xi^{(2)}} \int_{S_1} \left( \int_0^\infty e^{\alpha s} f_z^{(1)}(s) ds \right) \mathcal{P}(d(1, z)) + \\ + \int_{S_2} \left( \int_0^\infty e^{\alpha s} F_z^{(2)}(ds) \right) \mathcal{P}(d(2, z)) \leq \\ \leq k_\Gamma \mathbf{E} e^{\alpha \xi^{(2)}} \int_0^\infty \left( \alpha \int_0^\infty e^{\alpha s} \frac{1 - F^{(1)}(s+z)}{1 - F^{(1)}(z)} ds + 1 \right) \times \\ \times \frac{1 - F^{(1)}(z)}{T_1} dz + \\ + (1 - k_\Gamma) \int_0^\infty \left( \alpha \int_0^\infty e^{\alpha s} \frac{1 - F^{(2)}(s+z)}{1 - F^{(2)}(z)} ds + 1 \right) \frac{1 - F^{(2)}(z)}{T_2} dz \leq \\ \leq \frac{\alpha \mathbf{E} e^{\alpha \xi^{(2)}}}{T_1 + T_2} \int_0^\infty \int_0^\infty e^{\alpha s} C e^{-c(s+z)} ds dz + \\ + \frac{\alpha}{T_1 + T_2} \int_0^\infty \int_0^\infty e^{\alpha s} (1 - F^{(2)}(s+z)) ds dz + \\ + k_\Gamma \mathbf{E} e^{\alpha \xi^{(2)}} + (1 - k_\Gamma) \leq$$



$$\leq \frac{1}{T_1 + T_2} \left( \frac{\alpha C \mathbf{E} e^{\alpha \xi^{(2)}}}{(c - \alpha)c} + \alpha \widehat{M}_\alpha^{(2)} + T_1 \mathbf{E} e^{\alpha \xi^{(2)}} + T_2 \right) \stackrel{\text{def}}{=} \overline{M}_\alpha.$$

Поэтому

$$\begin{aligned} \int_{\mathcal{X}} K(\alpha, Z'_0, Z''_0) \mathcal{P}(dZ''_0) &= \\ &= \mathfrak{K}(\alpha, Z'_0) \leq \frac{C e^{\alpha R^*} M_\alpha(Z'_0) \overline{M}_\alpha}{(c - \alpha) (1 - \mathbf{E} e^{\alpha \xi^{(1)}} \mathbf{E} e^{\alpha \xi^{(2)}} q)}, \end{aligned}$$

что и завершает доказательство теоремы 1. ■

**Замечание 9.** Вообще говоря, условия теоремы 1 не позволяют оценить сверху величину  $\Theta$ . Если при исследовании конкретной технической системы такой оценки нет, то величину  $\Theta$  можно заменить величиной  $\Theta_1 \stackrel{\text{def}}{=} \frac{\mathbf{E} (\xi^{(1)} + \xi^{(2)})^2}{2(T_1 + T_2)}$ . Доказательство теоремы в этом случае меняется в части разделов 6.3.3–6.3.5: случай 6.3.3 вообще не учитывается, а вместо сл.в.  $\zeta_{\mathcal{F}_k}^*$  – остаточного времени пребывания процесса  $Z''_t$  в множестве  $\mathcal{S}_2$  – рассматривается сл.в.  $\zeta_{\mathcal{F}_k}^*$  – остаточное время периода регенерации процесса  $Z''_t$  – см. (11). В остальном доказательство не меняется. ▷

### Литература

1. БОРОВКОВ А.А. *Вероятностные процессы в теории массового обслуживания*. – М.: Наука, Гл. изд-во физ.-мат. литры, 1972. – 368 с.
2. ГНЕДЕНКО Б.В., БЕЛЯЕВ Ю.К., СОЛОВЬЕВ А.Д. *Математические методы в теории надежности*. – М.: Наука, 1965; также изд. 2, испр. и доп. – М.: URSS, 2013.
3. ГНЕДЕНКО Б.В., КОВАЛЕНКО И.Н. *Введение в теорию массового обслуживания*. – М.: Наука, 1966.
4. КЛИМОВ Г.П. *Теория вероятностей и математическая статистика*. – М.: Изд-во МГУ, 1983.

5. КОКС Д.Р., СМИТ В.Л. *Теория восстановления*. – М.: Сов. радио, 1967.
6. DAVIS M.H.A. *Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models* // J. Roy. Statist. Soc. Ser. B. – 1984. – Vol. 46, No. 3. – P. 353–388. – MR MR790622 (87g:60062).
7. GRIFFEATH D. *A maximal coupling for Markov chains* // Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete. – 1975. – Vol. 31, – Iss. 2. – P. 95–106.
8. LINDVALL T. *Lectures on the Coupling Method*. – Wiley, New York, 1992.
9. LORDEN G. *On Excess over the Boundary* // The Annals of Mathematical Statistics. – 1970. – Vol. 41, No. 2. – P. 520–527. – DOI: 10.1214/aoms/1177697092.
10. THORISSON H. *Coupling, Stationarity, and Regeneration*. – Springer, 2000.
11. VERETENNIKOV A.YU., BUTKOVSKY, O.A. *On asymptotics for Vaserstein coupling of Markov chains* // Stochastic Processes and their Applications. – 2013. – Vol. 123(9). P. 3518–3541.
12. VERETENNIKOV A. *Coupling method for Markov chains under integral Doeblin type condition* // Theory Stoch. Process. – 2002. – Vol. 8(24), No. 3–4. – P. 383–390.

## **ON EXPONENTIAL CONVERGENCE OF AVAILABILITY FACTOR**

**Galina Zverkina**, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand.Sc., senior researcher; Russian University of transport (MIIT), Moscow, associate professor (zverkina@gmail.com).

*Abstract: The availability factor is the probability that the system is working properly at a given moment of time. It's estimation and calculation are one the most important tasks in a reliability theory. The calculating of the stationary value of the availability factor is not difficult for the most cases. However, in real applications it is necessary to know how soon the value of the availability factor becomes sufficiently close to its limiting value, i.e. it is important to know the rate of convergence of the availability factor to its stationary value. In cases where the distributions function of the operating and recovery times are exponential, the question of the rate of convergence of the availability factor is the question of the rate of convergence of the solution of the Kolmogorov-Chapman equations with constant coefficients. This problem is solvable by means of Laplace transform. The most processes describing the behavior of reliability systems are regenerative, and for them the type of rate of convergence of the distribution to a stationary one is known - exponential or polynomial. The strong upper bounds of the rate of the convergence for regenerative processes can be obtained using the Lorden's inequality, when the distribution of the length of a regeneration period is known. But for reliability systems, it is possible to take into account the specific features of the investigated regeneration period. This article shows a method for constructing a strong exponential bounds for convergence rate of the availability factor for one restorable element.*

Keywords: availability factor, restorable element, bounds for convergence rate, exponential convergence.

УДК 519.2

ББК 22.171

DOI: 10.25728/ubs.2021.90.1

*Статья представлена к публикации  
членом редакционной коллегии В.М. Вишневым.*

*Поступила в редакцию 30.01.2021.*

*Дата опубликования 31.03.2021.*

## ОБ ОЦЕНКАХ СКОРОСТИ СХОДИМОСТИ ОДНОРОДНЫХ НЕЛИНЕЙНЫХ ЦЕПЕЙ МАРКОВА В ДИСКРЕТНОМ ВРЕМЕНИ<sup>1</sup>

Щеголев А. А.<sup>2</sup>

(Национальный исследовательский университет  
«Высшая школа экономики», Москва)

*Рассмотрена улучшенная оценка скорости сходимости однородных нелинейных марковских цепей в дискретном времени. Данный класс процессов нелинеен в терминах закона распределения, т.е. помимо зависимости от текущего состояния процесса переходные ядра также зависят и от вероятностного распределения в этот момент. Чаще всего такие процессы выступают в качестве предельных для больших систем зависимых цепей Маркова со взаимодействием. Полученная в работе оценка обобщает существующие результаты о сходимости с использованием переходных вероятностей за два шага. В частности показано, что данный подход не нарушает существования и единственности инвариантной меры при наложении условий, аналогичных использованным при построении оценки за один шаг. На примере нескольких нелинейных марковских цепей показано, что полученная оценка обладает более высокой скоростью сходимости, а также может быть использована в случаях, когда оценка за один шаг неприменима. Помимо этого, приведённые примеры иллюстрируют тот факт, что невыполнение условий сходимости для оценки за один шаг не препятствует сходимости некоторых однородных нелинейных цепей Маркова в дискретном времени.*

Ключевые слова: нелинейные марковские цепи, эргодичность, скорость сходимости.

### 1. Введение

Эргодические свойства классических цепей Маркова изучались в работах многих авторов, среди которых можно упомянуть А.А. Маркова, А.Н. Колмогорова, В. Дёблина, Дж.Л. Дуба, Р.Л. Добрушина. Существует несколько расширений теории марковских процессов, связанных с зависимостью процесса от его распределения. Так, в работе О. Оническу и Г. Михок [7]

---

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ, грант №20-01-00575.

<sup>2</sup> Александр Алексеевич Щеголев, аспирант НИУ ВШЭ (ashchegolev@hse.ru).

были предложены «цепи с полными связями» – процессы, зависящие от условного распределения на предыдущем шаге. Эргодические свойства данных процессов изучены школой румынских математиков, основные результаты которой изложены в монографии М. Иосифеску и С. Григореску [3]. Данное обобщение также известно и изучается в области символической динамики под названием  $g$ -мер [4]. Другое расширение связано с классом процессов, представленных Г.П. Маккином [6] в 1966 году. Данный тип процессов также изучался в работах ряда авторов, среди которых можно выделить монографии А.С. Шнитмана [8] и В.Н. Колокольцова [5]. Для рассматриваемого нами случая данное расширение получило название нелинейных марковских цепей. Здесь под нелинейностью подразумевается зависимость переходных функций процесса от его состояния и закона распределения в текущий момент. В отличие от «цепей с полными связями» [7], для данных процессов предполагается зависимость процесса от безусловного распределения процесса. Результаты исследования эргодических свойств однородных нелинейных цепей Маркова в дискретном времени были получены О.А. Бутковским [1]. В частности, было показано, что в нелинейном случае теории обычных цепей Маркова недостаточно, а также было установлено дополнительное условие для их сходимости. Такие процессы интересны тем, что чаще всего они выступают в качестве предельных для больших систем зависимых цепей Маркова со взаимодействием.

В данной работе сделано некоторое обобщение существующих результатов о сходимости для однородных нелинейных цепей Маркова с дискретным временем [1], используя оценку за несколько шагов. В первой части получено обобщение оценки эргодической сходимости, зависящей от переходных вероятностей за два шага. Во второй части представлены примеры однородных нелинейных цепей Маркова, для которых выполнены условия сходимости для нового результата, в то время как существующий результат за один шаг либо неприменим, либо приводит к более медленной сходимости.

## 2. Постановка задачи и основной результат

Условия сходимости и равномерной эргодичности для дискретных неприводимых апериодических однородных цепей Маркова с конечным пространством состояний показаны в различных классических источниках, и схожие результаты также существуют для более общих цепей Маркова (например, [9]).

Пусть задано измеримое пространство  $(E, \mathcal{E})$ , и  $\mathcal{P}(E)$  – множество вероятностных мер, определённых на этом пространстве. Тогда процесс  $(X_n^\mu)_{n \in \mathbb{Z}_+}$  – нелинейная цепь Маркова с пространством состояний  $(E, \mathcal{E})$ , начальным распределением  $\mu = \text{Law}(X_0^\mu)$ ,  $\mu \in \mathcal{P}(E)$ , и переходными вероятностями  $P_{\mu_n}(x, B) = \mathbb{P}_{\mu_n}(X_{n+1}^\mu \in B | X_n^\mu = x)$ , где  $x \in E$ ,  $B \in \mathcal{E}$ ,  $n \in \mathbb{Z}_+$  и  $\mu_n := \text{Law}(X_n^\mu)$ . Таким образом, переходное ядро зависит не только от состояния процесса в момент  $n$ , но и от его распределения в этот момент.

Пусть  $\mu, \nu \in \mathcal{P}(E)$ , тогда расстояние в метрике полной вариации между двумя вероятностными мерами может быть задано следующим образом:

$$\|\mu - \nu\|_{TV} = 2 \sup_{A \in \mathcal{E}} |\mu(A) - \nu(A)| = \int_E |\mu(dx) - \nu(dx)|.$$

Согласно результатам [1], нелинейная цепь Маркова является равномерно эргодическим процессом и существование единственной инвариантной меры  $\pi$  гарантировано, если она удовлетворяет следующим условиям:

$$(1) \quad \sup_{\mu, \nu \in \mathcal{P}(E)} \|P_\mu(x, \cdot) - P_\nu(y, \cdot)\|_{TV} \leq 2(1 - \alpha),$$

где  $0 < \alpha < 1$ ,  $x, y \in E$ ;

$$(2) \quad \|P_\mu(x, \cdot) - P_\nu(x, \cdot)\|_{TV} \leq \lambda \|\mu - \nu\|_{TV},$$

где  $\lambda \in [0, \alpha]$ ,  $x \in E$ ,  $\mu, \nu \in \mathcal{P}(E)$ .

Тогда сходимость является экспоненциальной, если  $\lambda < \alpha$  и

$$\|\mu_n - \pi\|_{TV} \leq 2(1 - (\alpha - \lambda))^n, \quad n \in \mathbb{Z}_+,$$

и в случае  $\lambda = \alpha$  имеем линейную сходимость

$$\|\mu_n - \pi\|_{TV} \leq \frac{2}{\lambda n}, \quad n \in \mathbb{Z}_+,$$

а для случая  $\lambda > \alpha$  в работе [1] были приведены контрпримеры, демонстрирующие как полное отсутствие, так и наличие нескольких инвариантных мер.

Цель данной работы заключается в том, чтобы показать, что эти контрпримеры относятся к определённому ограниченному классу нелинейных цепей, и для других нелинейных цепей Маркова условие  $\lambda > \alpha$  не препятствует экспоненциальной сходимости.

Обобщим результат [1] с использованием переходных ядер за два шага.

**Теорема 1** Существование и единственность инвариантной меры.

*Пусть процесс  $X$  имеет матрицу переходных вероятностей за два шага  $Q_\mu(x, A) := P_\mu(X_2 \in A | X_0 = x)$  и удовлетворяет следующим условиям:*

$$(3) \quad \sup_{\mu, \nu \in \mathcal{P}(E)} \|Q_\mu(x, \cdot) - Q_\nu(y, \cdot)\|_{TV} \leq 2(1 - \alpha_2),$$

где  $0 < \alpha_2 < 1$ ,  $x, y \in E$ ,

$$(4) \quad \|Q_\mu(x, \cdot) - Q_\nu(x, \cdot)\|_{TV} \leq \lambda_2 \|\mu - \nu\|_{TV},$$

где  $\lambda_2 \in [0, \alpha_2]$ ,  $x \in E$ ,  $\mu, \nu \in \mathcal{P}(E)$ ,

$$(5) \quad \|P_\mu(x, \cdot) - P_\nu(x, \cdot)\|_{TV} \leq \lambda_1 \|\mu - \nu\|_{TV}, \quad \lambda_1 < \infty.$$

*Тогда процесс  $X$  имеет единственную инвариантную меру  $\pi$  и для любой вероятностной меры  $\mu \in \mathcal{P}(E)$  справедлива сходимость:*

$$(6) \quad \|\mu_n - \pi\|_{TV} \leq \|\mu_0 - \pi\|_{TV} (1 - \alpha_2 + \lambda_2)^{\lfloor n/2 \rfloor} ((1 + \lambda_1) \mathbf{1}(n \text{ нечётное}) \vee 1),$$

и в случае  $\lambda_2 = \alpha_2$

$$(7) \quad \|\mu_n - \pi\|_{TV} \leq \frac{\|\mu_0 - \pi\|_{TV}}{1 + \frac{\lambda_2 n}{2} \|\mu_0 - \pi\|_{TV}} ((1 + \lambda_1) \mathbf{1}(n \text{ нечётное}) \vee 1).$$

Для доказательства данной теоремы потребуется вспомогательная теорема о сближении любых двух начальных вероятностных мер для данного процесса.

**Теорема 2.** Пусть процесс  $X$  имеет матрицу переходных вероятностей за два шага  $Q_\mu(x, B)$  и удовлетворяет условиям (3) и (4) теоремы 1. Тогда для любой пары вероятностных мер  $\mu, \nu \in \mathcal{P}(E)$  справедлива сходимость:

$$(8) \quad \begin{aligned} \|\mu_n - \nu_n\|_{TV} &\leq \\ &\leq \|\mu_0 - \nu_0\|_{TV} (1 - \alpha_2 + \lambda_2)^{\lfloor n/2 \rfloor} ((1 + \lambda_1) \mathbf{1}(n \text{ нечётное}) \vee 1), \end{aligned}$$

и в случае  $\lambda_2 = \alpha_2$

$$(9) \quad \|\mu_n - \nu_n\|_{TV} \leq \frac{\|\mu_0 - \nu_0\|_{TV}}{1 + \frac{\lambda_2 n}{2} \|\mu_0 - \nu_0\|_{TV}} ((1 + \lambda_1) \mathbf{1}(n \text{ нечётное}) \vee 1).$$

Доказательства данных теорем аналогичны [1] с использованием переходного ядра за два шага. Полное доказательство приведено с целью исправить вычислительные неточности доказательства [1], которые, однако, не влияли на конечный результат.

Докажем теорему 2.

**Доказательство.** Пусть даны  $P : E \times \mathcal{E} \rightarrow [0, 1]$  – переходное ядро, измеримая функция  $\varphi : E \rightarrow \mathbb{R}$  и вероятностная мера  $\mu \in \mathcal{P}(E)$ ; обозначим  $\mu P := \int_E P(x, dt) \mu(dx)$ ; в случае, когда  $P$  зависит от меры  $\mu$ , имеем:  $\mu_1(\mu) := \mu P_\mu := \int_E P_\mu(x, dt) \mu(dx)$ , тогда переходное ядро за два шага

$$Q_\mu(x, dy) = \int P_\mu(x, dx_1) P_{\mu_1(\mu)}(x_1, dy).$$

Рассмотрим расстояние в метрике полной вариации между мерами после применения переходного ядра за два шага. Для любых вероятностных мер  $\mu, \nu \in \mathcal{P}(E)$  обозначим

$$d\eta = ((d\mu/d\nu) \wedge 1) d\nu$$



и применим неравенство треугольника, в результате получаем

$$\begin{aligned} \|\mu Q_\mu - \nu Q_\nu\|_{TV} &= \|(\eta + (\mu - \eta))Q_\mu - (\eta - (\nu - \eta))Q_\nu\|_{TV} = \\ &= \int_E |\eta Q_\mu(dx) + (\mu - \eta)Q_\mu(dx) - \eta Q_\nu(dx) - (\nu - \eta)Q_\nu(dx)| \leq \\ &\leq \int_E |\eta Q_\mu(dx) - \eta Q_\nu(dx)| + \int_E |(\mu - \eta)Q_\mu(dx) - (\nu - \eta)Q_\nu(dx)| \leq \\ &\leq \|\eta Q_\mu - \eta Q_\nu\|_{TV} + \|(\mu - \eta)Q_\mu - (\nu - \eta)Q_\nu\|_{TV}. \end{aligned}$$

Рассмотрим первое слагаемое, применяя к нему неравенство Йенсена и (4), а также используя следующий факт:  $\eta(E) = 1 - \|\mu - \nu\|_{TV}/2$ . Получим

$$\begin{aligned} \|\eta Q_\mu - \eta Q_\nu\|_{TV} &= \int_E \left| \int_E Q_\mu(x, dy)\eta(dx) - \int_E Q_\nu(x, dy)\eta(dx) \right| \leq \\ &\leq \int_E \int_E |Q_\mu(x, dy) - Q_\nu(x, dy)|\eta(dx) \leq \\ &\leq \lambda_2 \|\mu - \nu\|_{TV} \left(1 - \frac{1}{2}\|\mu - \nu\|_{TV}\right). \end{aligned}$$

Тогда второе слагаемое

$$\begin{aligned} \|(\mu - \eta)Q_\mu - (\nu - \eta)Q_\nu\|_{TV} &= \int_E |(\mu - \eta)Q_\mu(dx) - (\nu - \eta)Q_\nu(dx)| = \\ &= \int_E \left| \int_E Q_\mu(x, dy)(\mu - \eta)(dx) - \int_E Q_\nu(x', dy)(\nu - \eta)(dx') \right|. \end{aligned}$$

Напомним, что  $\mu_n = \text{Law}(X_n^\mu)$ ,  $\nu_n = \text{Law}(X_n^\nu)$ , обозначим  $p_0 = \|\mu_0 - \nu_0\|_{TV}/2$ , предполагая  $p_0 > 0$  (если  $p_0 = 0$ , тогда  $p_2 = 0$  и т.д.).

Оценим выражение  $\|\mu_2 - \nu_2\|_{TV}$  сверху.

$$\begin{aligned} \|\mu_2 - \nu_2\|_{TV} &\leq \lambda_2 \|\mu_0 - \nu_0\|_{TV} \left(1 - \frac{1}{2}\|\mu_0 - \nu_0\|_{TV}\right) + \\ &+ p_0 \int \left| \int Q_\mu(x, dy) \frac{(\mu_0 - \eta_0)(dx)}{p_0} - \int Q_\nu(x', dy) \frac{(\nu_0 - \eta_0)(dx')}{p_0} \right| = \\ &= 2p_0 \lambda_2 (1 - p_0) + p_0 \int \left| \iint (Q_\mu(x, dy) - Q_\nu(x', dy)) \frac{(\mu_0 - \eta_0)(dx)}{p_0} \frac{(\nu_0 - \eta_0)(dx')}{p_0} \right| \leq \end{aligned}$$

$$\begin{aligned} &\leq 2p_0\lambda_2(1-p_0) + p_0 \iiint |Q_\mu(x, dy) - Q_\nu(x', dy)| \frac{(\mu_0 - \eta_0)(dx)}{p_0} \frac{(\nu_0 - \eta_0)(dx')}{p_0} \leq \\ &\leq 2p_0\lambda_2(1-p_0) + 2(1-\alpha_2)p_0 \iint \frac{(\mu_0 - \eta_0)(dx)}{p_0} \frac{(\nu_0 - \eta_0)(dx')}{p_0} = \\ &= 2p_0\lambda_2(1-p_0) + 2p_0(1-\alpha_2) = 2p_0(\lambda_2 - \lambda_2p_0 + 1 - \alpha_2). \end{aligned}$$

Если  $\lambda_2 < \alpha_2$ , получаем

$$\|\mu_2 - \nu_2\|_{TV} \leq \|\mu_0 - \nu_0\|_{TV}(1 - \alpha_2 + \lambda_2),$$

в то время как в случае  $\lambda_2 = \alpha_2$  имеем

$$\|\mu_2 - \nu_2\|_{TV} \leq 2p_0(1 - \lambda_2p_0),$$

или

$$p_2 \leq p_0(1 - \lambda_2p_0).$$

Для случая  $2n + 1$  имеем:

$$\begin{aligned} &\|\mu_{2n+1} - \nu_{2n+1}\|_{TV} \leq \lambda_1 \|\mu_{2n} - \nu_{2n}\|_{TV} \left(1 - \frac{1}{2} \|\mu_{2n} - \nu_{2n}\|_{TV}\right) + \\ &+ p_{2n} \int \left| \int P_\mu(x, dy) \frac{(\mu_{2n} - \eta_{2n})(dx)}{p_{2n}} - \int P_\nu(x', dy) \frac{(\nu_{2n} - \eta_{2n})(dx')}{p_{2n}} \right| = \\ = 2p_{2n} \lambda_1(1 - p_{2n}) + p_{2n} \int \left| \iint (P_\mu(x, dy) - P_\nu(x', dy)) \frac{(\mu_{2n} - \eta_{2n})(dx)}{p_{2n}} \frac{(\nu_{2n} - \eta_{2n})(dx')}{p_{2n}} \right| \leq \\ &\leq 2p_{2n} \lambda_1(1 - p_{2n}) + p_{2n} \iiint |P_\mu(x, dy) - P_\nu(x', dy)| \frac{(\mu_{2n} - \eta_{2n})(dx)}{p_{2n}} \frac{(\nu_{2n} - \eta_{2n})(dx')}{p_{2n}} \leq \\ &\leq 2p_{2n}(\lambda_1(1 - p_{2n}) + 1) = (1 + \lambda_1) \|\mu_{2n} - \nu_{2n}\|_{TV}. \end{aligned}$$

Следовательно,

$$(10) \quad \|\mu_{2n+1} - \nu_{2n+1}\|_{TV} \leq (1 + \lambda_1) \|\mu_{2n} - \nu_{2n}\|_{TV}.$$

Таким образом, итерируя оценку для  $\lambda_2 < \alpha_2$ , получаем по индукции

$$\|\mu_n - \nu_n\|_{TV} \leq \|\mu_0 - \nu_0\|_{TV}(1 - \alpha_2 + \lambda_2)^{\lceil n/2 \rceil} ((1 + \lambda_1) \mathbf{1}(n \text{ нечётное}) \vee 1).$$

В случае  $\alpha_2 = \lambda_2$  применим следующую лемму.

**Лемма 1.** Пусть  $a_0, a_1, \dots$  – некоторая последовательность положительных чисел. Предположим, что  $0 < a_0 \leq 1$  и справедлива оценка

$$a_{n+1} \leq a_n(1 - \psi(a_n)), \quad n \in \mathbb{Z}_+,$$

где  $\psi : [0, \infty) \rightarrow [0, 1]$  – непрерывная неубывающая функция с  $\psi(0) = 0$  и  $\psi(x) > 0$  при  $x > 0$ . Тогда

$$a_n \leq g^{-1}(n)$$

для всех  $n \in \mathbb{Z}_+$ , где

$$g(x) = \int_x^{a_0} \frac{dt}{t\psi(t)}, \quad 0 < x \leq 1.$$

Данная лемма в несколько ином варианте и ее доказательство приведены в [2]. Поскольку здесь используется слегка измененный вариант с другим верхним пределом в интеграле, доказательство приводится ради удобства читателя, несмотря на то, что оно совпадает с первоисточником.

**Доказательство.** Заметим, что функция  $g^{-1}$  существует, поскольку  $g$  неограничена, неотрицательна и строго убывает. Тогда из неотрицательности  $\psi$  следует, что  $a_{n+1} \leq a_n$  для любого  $n \in \mathbb{N}$ . Тогда найдется  $s \in [a_{n+1}, a_n]$  такое, что

$$g(a_{n+1}) - g(a_n) = g'(s)(a_{n+1} - a_n) = -\frac{a_{n+1} - a_n}{s\psi(s)} \geq \frac{a_n\psi(a_n)}{s\psi(s)} \geq 1.$$

Таким образом,  $g(a_n) \geq n$  и  $a_n \leq g^{-1}(n)$ .

Применяя лемму 1 для  $a_{n+2}$  и  $\psi(t) = \lambda_2 t$  получаем

$$p_2 \leq g^{-1}(n) = \frac{1}{\lambda_2 n + \frac{1}{p_0}} = \frac{p_0}{1 + p_0 \lambda_2 n}.$$

Обобщая результат для нечётных  $n$ , используя (10), имеем:

$$\|\mu_n - \nu_n\|_{TV} \leq \frac{\|\mu_0 - \nu_0\|_{TV}}{1 + \frac{\lambda_2 n}{2} \|\mu_0 - \nu_0\|_{TV}} ((1 + \lambda_1) \mathbf{1}(n \text{ нечётное}) \vee 1).$$

Далее перейдём к доказательству теоремы 1.

**Доказательство.** Рассмотрим последовательность вероятностных мер  $(\mu_n)_{n \in \mathbb{N}}$ . Согласно теореме 2, в силу (8) и (9), для любых  $m, n \in \mathbb{N}$

$$\|\mu_n - \mu_{n+m}\|_{TV} \leq \frac{\|\mu_0 - \mu_m\|_{TV}}{1 + \frac{\lambda_2 n}{2} \|\mu_0 - \mu_m\|_{TV}} ((1 + \lambda_1) \mathbf{1}(n \text{ нечётное}) \vee 1).$$

Тогда  $(\mu_n)_{n \in \mathbb{N}}$  является последовательностью Коши в полном метрическом пространстве  $(\mathcal{P}(E), \|\cdot\|_{TV})$  и можно найти  $\pi \in \mathcal{P}(E)$ , такое что  $\lim_{n \rightarrow \infty} \|\mu_n - \pi\|_{TV} = 0$ .

Покажем, что предельная мера  $\pi$  инвариантна. Для этого используем неравенство треугольника и условие (10) при  $n \rightarrow \infty$ :

$$\|\pi P_\pi - \mu_{n+1}\|_{TV} = \|\pi P_\pi - \mu_n P_{\mu_n}\|_{TV} \leq (1 + \lambda_1) \|\pi - \mu_n\|_{TV} \rightarrow 0,$$

тогда как  $\mu_{n+1} \rightarrow \pi$ , имеем

$$\|\pi P_\pi - \pi\|_{TV} \leq \|\pi P_\pi - \mu_{n+1}\|_{TV} + \|\mu_{n+1} - \pi\|_{TV} \rightarrow 0.$$

Откуда получаем  $\pi = \pi P_\pi$ .

Для доказательства единственности инвариантной меры  $\pi$  предположим, что  $\nu \in \mathcal{P}(E)$  такая, что  $\nu \neq \pi$  и  $\nu = \nu P_\nu$ , тогда итеративно применяя результаты (8) и (9) получаем противоречие

$$\|\nu - \pi\|_{TV} = \|\nu Q_\nu - \pi Q_\pi\|_{TV} < \|\nu - \pi\|_{TV}.$$

Таким образом, процесс  $(X_n^\mu)_{n \in \mathbb{Z}}$  имеет единственную инвариантную меру  $\pi$ .

В разделе 3 мы проиллюстрируем, что данная оценка может быть лучше, чем оценка для одношагового перехода.

### 3. Примеры

#### 3.1. ПРИМЕР НЕЛИНЕЙНОЙ ЦЕПИ МАРКОВА, $\alpha > 0, \lambda > \alpha$

Рассмотрим однородную нелинейную цепь Маркова в дискретном времени  $X_n^\mu$  с пространством состояний  $(E, \mathcal{E}) = (\{1, 2, 3, 4\}, 2^{\{1,2,3,4\}})$ , начальным распределением  $\mu_0$  и матрицей переходных вероятностей  $P_{\mu_0}(i, j)$ , заданной следующим образом:

$$\mu_0 = (\mu(\{1\}) \quad \mu(\{2\}) \quad \mu(\{3\}) \quad \mu(\{4\}))$$

$$P_{\mu_0}(i, j) = \begin{pmatrix} 0,001 + \gamma\mu(\{1\}) & 0,001 + \gamma\mu(\{1\}) & 0,499 - \gamma\mu(\{1\}) & 0,499 - \gamma\mu(\{1\}) \\ 0,499 & 0,499 & 0,001 & 0,001 \\ 0,499 & 0,001 & 0,499 & 0,001 \\ 0,001 & 0,499 & 0,001 & 0,499 \end{pmatrix},$$

где  $0 < \gamma < 0,25$ .

Можем заметить, что для заданного процесса условия (1) и (2) гарантируют сходимость к инвариантной мере только в случае  $\gamma \leq 0,004$ , поскольку  $\alpha = 0,004$  и  $\lambda = \gamma$ .

Оценим соответствующую матрицу переходных вероятностей за два шага  $P_{\mu_2}(i, j)$ . Обозначим в элементах матрицы  $\mu_i = \mu(\{i\})$  для краткости, а также  $\xi(\mu, \gamma) = (\mu_1(\gamma\mu_1 +$

$+0,001) + 0,499\mu_2 + 0,499\mu_3 + 0,001\mu_4$ ) и  $\zeta(\mu, \gamma, x) = (\gamma\mu_1 + +0,001)(\gamma\xi(\mu, \gamma) + x)$  и  $\delta(\gamma, \mu, x, y) = y\gamma\mu_1 + \zeta(\gamma, \mu, x)$ , тогда переходное ядро за два шага имеет вид

$$P_{\mu_2}(i, j) = \begin{pmatrix} \delta(\gamma, \mu, 0,001, -0,001) + 0,249999 & \delta(\gamma, \mu, 0,001, -0,001) + 0,249999 & \dots \\ 0,499\gamma\xi(\mu, \gamma) + 0,25 & 0,499\gamma\xi(\mu, \gamma) + 0,25 & \dots \\ 0,499\gamma\xi(\mu, \gamma) + 0,25 & 0,499\gamma\xi(\mu, \gamma) + 0,001996 & \dots \\ 0,001\gamma\xi(\mu, \gamma) + 0,25 & 0,001\gamma\xi(\mu, \gamma) + 0,498004 & \dots \\ \dots & \delta(\gamma, \mu, -0,499, -0,499) + 0,249501 & \delta(\gamma, \mu, -0,499, -0,499) + 0,249501 \\ \dots & -0,499\gamma\xi(\mu, \gamma) + 0,25 & -0,499\gamma\xi(\mu, \gamma) + 0,25 \\ \dots & -0,499\gamma\xi(\mu, \gamma) + 0,498004 & -0,499\gamma\xi(\mu, \gamma) + 0,25 \\ \dots & -0,001\gamma\xi(\mu, \gamma) + 0,001996 & -0,001\gamma\xi(\mu, \gamma) + 0,25 \end{pmatrix}$$

В этом случае  $\alpha_2 = 0,503992$ , а  $\lambda_2 \leq \gamma$ . Потому имеем  $\lambda_2 < \alpha_2$ , что приводит к экспоненциальной сходимости процесса.

### 3.2. ПРИМЕР НЕЛИНЕЙНОЙ ЦЕПИ МАРКОВА, $\alpha = 0, \lambda > \alpha$

Покажем, что данный результат может быть использован в случаях, когда оценка за один шаг неприменима, и, в то же время, невыполнение условий [1] не препятствует экспоненциальной сходимости для некоторых нелинейных марковских цепей. Рассмотрим следующую дискретную нелинейную Марковскую цепь  $X_n^\mu$  с пространством состояний  $(E, \mathcal{E}) = (\{1, 2, 3, 4\}, 2^{\{1,2,3,4\}})$ , начальным распределением  $\mu_0$  и матрицей переходных вероятностей  $P_{\mu_0}(i, j)$ , заданной следующим образом:

$$\mu_0 = (\mu(\{1\}) \quad \mu(\{2\}) \quad \mu(\{3\}) \quad \mu(\{4\})),$$

$$P_{\mu_0}(i, j) = \begin{pmatrix} 0 & \gamma\mu(\{1\}) & 0,5 - \gamma\mu(\{1\}) & 0,5 \\ 0,5 & 0,5 & 0 & 0 \\ 0,5 & 0 & 0,5 & 0 \\ 0 & 0,5 & 0 & 0,5 \end{pmatrix},$$

где  $0 < \gamma < 0,5$ .

Можем заметить, что для заданного процесса условия (1) и (2) не гарантируют сходимости к инвариантной мере, так как  $\lambda > \alpha$ , поскольку  $\alpha = 0$  и  $\lambda = \gamma$ .

Однако, если рассмотреть соответствующую матрицу переходных вероятностей за два шага  $Q_{\mu_0}(i, j)$ ,

$$Q_{\mu_0}(i, j) = \begin{pmatrix} 0,25 & 0,5\gamma\mu_1 + 0,25 & -0,5\gamma\mu_1 + 0,25 & 0,25 \\ 0,25 & 0,25\gamma(\mu_2 + \mu_3) + 0,25 & -0,25\gamma(\mu_2 + \mu_3) + 0,25 & 0,25 \\ 0,25 & 0,25\gamma(\mu_2 + \mu_3) & -0,25\gamma(\mu_2 + \mu_3) + 0,5 & 0,25 \\ 0,25 & 0,5 & 0 & 0,25 \end{pmatrix},$$

можно получить иной результат. Имеем  $\lambda_2 < \alpha_2$ , поскольку  $\lambda_2 = \gamma/2$ , в то время как  $\alpha_2$  достигает минимального значения для пары состояний  $\{3, 4\}$  со значением в интервале  $[0,5; 0,5 + 0,25\gamma]$ . Таким образом, предлагаемая оценка может гарантировать экспоненциальную сходимость в некоторых случаях, когда существующий результат [1] не работает.

#### 4. Заключение

В статье предложена улучшенная оценка скорости сходимости однородных нелинейных цепей Маркова с дискретным временем путём обобщения существующих результатов о сходимости и получения оценки за несколько шагов. Эта оценка приводит к лучшей сходимости и даже может быть применима в случаях, когда оценка, сделанная за один шаг, не может гарантировать сходимости. В последнем разделе работы приводится пример однородной нелинейной цепи Маркова с конечным пространством состояний и дискретным временем, которая иллюстрирует этот результат. Кроме того, этот пример показывает, что невыполнение условий сходимости, предложенных в [1], не препятствует существованию единственной инвариантной меры и экспоненциальной сходимости для однородных нелинейных цепей Маркова в дискретном времени.

#### Литература

1. БУТКОВСКИЙ О.А. *Об эргодических свойствах нелинейных марковских цепей и стохастических уравнений Макакина–Власова* // Теория вероятностей и ее применения. – 2013. – Т. 58, №4. – С. 782–794.
2. БУТКОВСКИЙ О.А. *Предельные теоремы для марковских процессов* : Дис. – Московский государственный университет им. М.В. Ломоносова (МГУ). Механико-математический факультет, 2013.

3. IOSIFESCU M., GRIGORESCU S. *Dependence with complete connections and its applications* // Cambridge Tracts in Mathematics. – Cambridge University Press, Cambridge, 1990. – Vol. 96.
4. KEANE M. *Strongly mixing g-measures* // Inventiones Mathematicae. – 1972. – Vol. 16, No. 4. – P. 309–324.
5. KOLOKOLTSOV V.N. *Nonlinear Markov processes and kinetic equations* // Cambridge Tracts in Mathematics. – Cambridge University Press, Cambridge, 2010. – Vol. 182.
6. MCKEAN H.P. *A class of Markov processes associated with nonlinear parabolic equations* // Proc. of the National Academy of Sciences of the United States of America. – 1966. – Vol. 56. – P. 1907–1911.
7. ONICESCU O., MIHOC G. *Sur les chaînes de variables statistiques.* // Bull. Sci. Math. – 1935. – Vol. 59, No. 2. – P. 174–192.
8. SZNITMAN A.-S. *Topics in propagation of chaos* // École d'Été de Probabilités de Saint-Flour XIX.–1989. – Lecture Notes in Math. – Springer, Berlin, 1991. – Vol. 1464. – P. 165–251.
9. VERETENNIKOV A.YU. *Ergodic Markov processes and Poisson equations (lecture notes)* // Modern problems of stochastic analysis and statistics. – 2017. – Vol. 208. – P. 457–511.

## **ON RATE OF CONVERGENCE ESTIMATES FOR HOMOGENEOUS DISCRETE-TIME NONLINEAR MARKOV CHAINS**

**Aleksandr Shchegolev**, National Research University Higher School of Economics, Moscow, postgraduate student (ashchegolev@hse.ru).

*Abstract: The paper studies an improved estimate for the rate of convergence for nonlinear homogeneous discrete-time Markov chains. These processes are nonlinear in terms of the distribution law. Hence, the transition kernels are dependent on the current state. Such processes often act as limits for large-scale systems of dependent Markov chains with interaction. The paper generalizes the convergence results by taking the estimate over two steps. Such an approach keeps the existence and uniqueness results under assumptions that are analogical to the one-step result. It is shown that such an approach may lead to a better rate of convergence. Several examples provided illustrating the fact that the suggested estimate may have a better rate of convergence than the original one. Also, it is shown that the new estimate may even be applicable in some cases when the conditions of the result on one step cannot guarantee any convergence. Finally, these examples depict that the original conditions may not be an obstacle for the convergence of nonlinear Markov chains.*

Keywords: nonlinear Markov chains, ergodicity, rate of convergence.

УДК 519.2

ББК 22.171

DOI: 10.25728/ubs.2021.90.2

*Статья представлена к публикации  
членом редакционной коллегии Э.Ю. Калимулиной.*

*Поступила в редакцию 19.03.2021.*

*Дата опубликования 31.03.2021.*



# ИССЛЕДОВАНИЕ ПОЛНЫХ ОДНОРОДНЫХ РЕСУРСНЫХ СЕТЕЙ С «ЖАДНЫМИ» ВЕРШИНАМИ: ЗОНА «ДОСТАТОЧНОГО БОЛЬШОГО» РЕСУРСА

Чаплинская Н. В.<sup>1</sup>

(ФГБУН Институт проблем управления  
им. В.А. Трапезникова РАН, Москва)

*«Ресурсная сеть с жадными вершинами» – модификация графовой динамической модели «ресурсная сеть». На каждом такте дискретного времени вершины графа передают друг другу ресурс по ребрам с ограниченной пропускной способностью, причем сначала передают имеющийся ресурс себе в петлю, а затем оставшийся ресурс, если он имеется, распределяют по исходящим ребрам по правилам «стандартной» ресурсной сети. Это два правила с пороговым переключением: если вершина имеет ресурс, превышающий суммарную пропускную способность всех ее исходящих ребер, она передает в каждое исходящее ребро полную его пропускную способность, в противном случае отдает весь свой ресурс, распределяя его пропорционально пропускным способностям исходящих ребер. Рассмотрен частный случай: полная однородная ресурсная сеть с «жадными» вершинами. Для сети такого вида существуют два пороговых значения суммарного ресурса, разделяющих зоны различного поведения сети: первое разделяет зоны «недостаточного» и «достаточного» ресурса, второе – зоны «достаточного малого» и «достаточного большого» ресурса. В данной статье исследована последняя зона – зона «достаточного большого» ресурса: описан процесс функционирования сети, найдено предельное состояние сети.*

Ключевые слова: графовая динамическая пороговая модель, ресурсная сеть, модель «жадных» вершин, зона «достаточного большого» ресурса.

## 1. Введение

Ресурсной сетью называют графовую динамическую потоковую модель замкнутой системы, в которой ресурс в дискретном времени перераспределяется между вершинами, не поступая извне и не расходуясь. Эта модель была предложена впервые в 2009 году в [13], а свое дальнейшее развитие получила в [10, 12]. Ресурсная сеть описывается ориентированным взвешенным графом. Веса ребер называются пропускными способностями. Модель характеризуется двумя правилами распределения

<sup>1</sup> Надежда Васильевна Чаплинская, техник (nadya1462@gmail.com).

ресурса между вершинами. Выбор правила определяется пороговым переключением, которое происходит в зависимости от количества ресурса в вершине и от пропускных способностей ребер, выходящих из этой вершины.

В [10] для данной модели были получены результаты, позволяющие для сети любой топологии с любым значением суммарного ресурса и любым его начальным распределением между вершинами определить предельный поток и предельное состояние сети. В [10] также была определена формула, позволяющая для такой произвольной сети найти пороговое значение суммарного ресурса, разделяющее зоны различного поведения сети.

При *малом ресурсе*, т.е. при суммарном ресурсе, не превосходящем порогового значения, динамика сети описывается моделью рассеяния на графах [15, 16]. Для графов, имеющих постоянную топологию, модели рассеяния, в свою очередь, описываются однородными цепями Маркова, которые широко используются при моделировании процессов из различных предметных областей, например, при решении задач о нахождении консенсуса в многоагентных системах [5, 6], решении задач о нахождении количественной оценки различных видов риска и их влияния на финансовое состояние предприятия [9], в социальных сетях [7] и т.д.

При *большом ресурсе*, т.е. при суммарном ресурсе выше порогового значения, динамика сети описывается неоднородной цепью Маркова. В [10] для случая большого ресурса были обнаружены вершины, накапливающие излишки ресурса, – вершины-аттракторы.

С момента появления ресурсной сети в 2009 году в [13], были рассмотрены некоторые модификации этой модели. Была представлена модель с ограничениями на ёмкость вершин (монография [12]). Были предложены и описаны ресурсные сети на графах с нестандартной достижимостью (см. [8]): в работе [1] – с вентильной достижимостью, в работе [2] – с магнитной достижимостью. В [3] были введены двухресурсные сети и рассмотрена топология полных однородных несимметричных двусторонних

ресурсных сетей с петлями. В продолжение этой работы в [4] были рассмотрены двухресурсные сети с магнитной достижимостью. В [14] были описаны динамические ресурсные сети.

Настоящая статья является продолжением работы [11], в которой была рассмотрена еще одна модификация ресурсных сетей – ресурсная сеть с «жадными» вершинами. Отличие этой модели заключается в том, что каждая вершина передает ресурс сначала в свою петлю (если она есть) и только затем остатки распределяет в другие вершины. Рассматривается частный случай: полная однородная сеть с петлями. Для этого случая были найдены два пороговых значения суммарного ресурса, разделяющих зоны различного поведения сети: первое разделяет зоны «недостаточного» и «достаточного» ресурса, второе – зоны «достаточного малого» и «достаточного большого» ресурса. В [11] были исследованы зоны «недостаточного» и «достаточного малого» ресурса. В настоящей статье исследуется зона «достаточного большого» ресурса.

## 2. Ресурсные сети – основные понятия

Ресурсная сеть с «жадными» вершинами есть модификация обычной ресурсной сети, наследующая все ее базовые понятия. Приведем основные определения «стандартной» ресурсной сети, а также опишем процесс ее функционирования.

*Ресурсная сеть* – это связный ориентированный граф  $G = (V, E)$  с вершинами  $v_i \in V$ ,  $|V| = n$ , ребрам  $e_{ij} = (v_i, v_j) \in E$  которого приписаны неотрицательные веса  $r_{ij}$ , обозначающие *пропускные способности* этих ребер. Матрицу  $R = (r_{ij})_{n \times n}$  называют *матрицей пропускных способностей*.

Вершинам графа приписаны *ресурсы*  $q_i(t)$  – неотрицательные числа, изменяющиеся в дискретном времени  $t$ . Каждая вершина может хранить неограниченное количество ресурса. Вектор  $Q(t) = (q_1(t), \dots, q_n(t))$ , элементами которого являются значения ресурсов в вершинах в момент времени  $t$ , называется *состоянием*  $Q(t)$  сети в момент времени  $t$ .

Состояние  $Q^* = (q_1^*, \dots, q_n^*)$  называется *предельным*, если для

любого  $\varepsilon > 0$  существует  $t_\varepsilon$  такое, что для всех  $t > t_\varepsilon$   $|q_i^* - q_i(t)| < \varepsilon$ ,  $i = 1, 2, \dots, n$ .

В сети выполняется закон сохранения ресурса:  $\forall t \sum_{i=1}^n q_i(t) = W$ , где  $W$  – это суммарный ресурс всех вершин сети.

$r_i^{in} = \sum_{j=1}^n r_{ji}$  и  $r_i^{out} = \sum_{j=1}^n r_{ij}$  – входная и выходная пропускные способности вершины соответственно.

В каждый момент времени все вершины сети  $v_i, i = 1, \dots, n$ , передают имеющийся в них ресурс по выходящим ребрам. Правила передачи ресурса в сети *стандартной модели* следующие: в момент  $t$  вершина  $v_i$  отдает в ребро  $e_{ij}$ , соединяющее ее с вершиной  $v_j$

- $r_{ij}$  единиц ресурса, если  $q_i(t) > r_i^{out}$  (правило 1);
- $\frac{r_{ij}}{r_i^{out}} q_i(t)$  единиц ресурса, если  $q_i(t) \leq r_i^{out}$  (правило 2).

Таким образом, правило 1 применяется тогда, когда вершина содержит ресурса больше, чем может отдать по исходящим ребрам. В этом случае по каждому ребру отдается ресурс, равный пропускной способности этого ребра. Всего вершина отдаст  $r_i^{out} = \sum_{j=1}^n r_{ij}$  ресурса. Если условие применения правила 1 не выполняется, т.е. если в вершине недостаточно ресурса для передачи полной пропускной способности по каждому ребру, то работает правило 2. В этом случае весь свой имеющийся ресурс вершина распределяет пропорционально пропускным способностям ребер. При  $q_i(t) = r_i^{out}$  применение правил 1 и 2 дает одинаковые результаты.

Для каждого момента времени  $t$  множество вершин, функционирующих по правилу 1, будем называть *зоной*  $Z^+(t)$ , по правилу 2 – *зоной*  $Z^-(t)$ .

Вершины, имеющие ресурс, отдают его одновременно на каждом такте. Ресурс, выходящий из вершины  $v_i$  по ребру  $e_{ij}$  в момент  $t$ , приходит в вершину  $v_j$  в момент  $t + 1$ . Между моментами  $t$  и  $t + 1$  ресурс находится в ребре  $e_{ij}$ . Этот ресурс называется

ся потоком  $f_{ij}(t)$ . Величина  $f_i^{out}(t) = \sum_{j=1}^n f_{ij}(t)$  – ресурс, выходящий из вершины  $v_i$  в момент  $t$ . Величина  $f_j^{in}(t+1) = \sum_{i=1}^n f_{ij}(t)$  – ресурс, входящий в вершину  $v_j$  в момент  $t+1$ .

Пороговое значение ресурса  $T$  – важная характеристика «стандартной» ресурсной сети.  $T$  разделяет зоны различного поведения сети: при  $W \leq T$ , начиная с некоторого момента времени  $t'$ , все вершины переходят в зону  $Z^-(t)$ ; при  $W > T$ , начиная с некоторого момента времени  $t''$ , зона  $Z^+(t)$  не будет пуста и станет постоянной по мощности. В [10] доказано, что в любой сети, кроме поглощающей, пороговое значение ресурса  $T$  существует и единственно. В [10] также предложены формулы, которые позволяют определить значение  $T$  для любого класса сетей (кроме поглощающей).

### 3. Модель сети с «жадными» вершинами

#### 3.1. ОПИСАНИЕ И ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Модель сети с «жадными» вершинами была предложена и формально описана в [11]. Напомним основные понятия и особенности поведения данной модели.

Рассматривается ресурсная сеть с наличием хотя бы одной петли. Ее функционирование определяется следующим образом:

Вершина  $v_i$ , имеющая петлю, на такте  $t$  отдает:

- $r_{ii}$  ресурса в петлю, если  $q_i(t) > r_{ii}$ , и оставшийся ресурс  $\Delta q_i(t) = q_i(t) - r_{ii}$  распределяет между другими вершинами по правилам стандартной ресурсной сети,

- $q_i(t)$  ресурса в петлю, если  $q_i(t) \leq r_{ii}$  (весь имеющийся ресурс уходит в петлю).

Вершина, не имеющая петлю, отдает ресурс другим вершинам по правилам стандартной ресурсной сети.

Другими словами, все вершины с петлями являются «запасливыми» (или «жадными»): они передают имеющийся ресурс в первую очередь себе в петлю и уже затем распределяют остатки (если они есть) между другими вершинами.

Будем говорить, что сеть с «жадными» вершинами *остановилась* на такте  $t$ , если  $\forall i = \overline{1, n} \ q_i(t) \leq r_{ii}$ . Это означает, что если количество ресурса во всех вершинах не превосходит пропускных способностей петель, то ресурс не перераспределяется между вершинами. Остановка сети является частным случаем ее стабилизации.

Вершина  $v_i$  называется *насыщенной* в момент времени  $t$ , если  $q_i(t) \geq r_{ii}$ , и *ненасыщенной*, если  $q_i(t) < r_{ii}$ . В [11] доказано, что вершина, насыщенная в момент времени  $t$ , в момент времени  $t + 1$  также останется насыщенной.

Ограничение на топологию сетей: исследуются полные однородные ресурсные сети с «жадными» вершинами. Ресурсная сеть называется *полной*, если  $r_{ij} \neq 0 \ \forall i, j = \overline{1, n}$ . *Однородность* сети означает, что  $r_{ij} = r \ \forall i, j = \overline{1, n}$ .

### 3.2. ПОРОГОВЫЕ ЗНАЧЕНИЯ РЕСУРСА

В [11] было показано, что для полных однородных ресурсных сетей с «жадными» вершинами существуют два пороговых значения суммарного ресурса  $T_1 = rn$  и  $T_2 = rn^2$ , разделяющих зоны различного поведения сети (рис. 1).  $T_1$  разделяет зоны «недостаточного» и «достаточного» ресурса. Зона «достаточного» ресурса в свою очередь делится порогом  $T_2$  на зоны «достаточного малого» и «достаточного большого» ресурса.

Для зоны «недостаточного» ресурса было доказано, что сеть останавливается (за конечное число тактов или асимптотически, в зависимости от начального состояния сети).

В зоне «достаточного» ресурса сеть никогда не останавливается. При этом за конечное число тактов все вершины сети становятся насыщенными, и с момента их насыщения сеть функционирует эквивалентно стандартной полной однородной ресурсной сети без петель.

«Достаточный» ресурс, по аналогии со стандартной сетью, делится порогом  $T_2 = rn^2$  на «малый» ( $T_1 < W \leq T_2$ ) и «большой» ресурс ( $W > T_2$ ). Для зоны «достаточного малого» ресурса была доказана стабилизация сети и найдено ее предельное состояние.

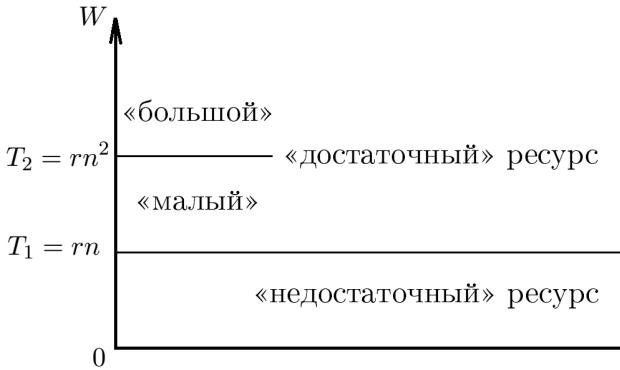


Рис. 1. Схема пороговых значений суммарного ресурса в сети с «жадными» вершинами

### 3.3. «ДОСТАТОЧНЫЙ БОЛЬШОЙ» РЕСУРС

Рассмотрим функционирование сети при  $W > T_2 = rn^2$ .

«Большой ресурс» означает, что при любом начальном состоянии  $Q(0)$  в любой момент времени зона  $Z^+(t)$  будет непуста.

**Утверждение 1.** При большом суммарном ресурсе  $W > T_2$  ненасыщенные вершины насыщаются за 1 такт.

**Доказательство.** В сети есть хотя бы одна вершина из зоны  $Z^+$ , а, значит, ненасыщенная вершина получит ресурса не меньше  $r$ . Таким образом, на такте  $t + 1$  она будет иметь ресурс размер не меньше  $r$ , т.е. станет насыщенной.

Учитывая утверждение 1, далее будем рассматривать сеть с уже насыщенными вершинами.

Количество вершин из зоны  $Z^+(t)$  обозначим  $k_t$ . Пусть на такте  $t = 0$  вершины пронумерованы по убыванию имеющегося в них ресурса. То есть в зоне  $Z^+(0)$  находятся вершины  $v_i$ ,  $i = \overline{1, k_0}$ , в зоне  $Z^-(0)$  – вершины  $v_i$ ,  $i = \overline{k_0 + 1, n}$ :

$$q_1(0) \geq \dots \geq q_{k_0}(0) > rn \geq q_{k_0+1}(0) \geq q_n(0).$$

Введем понятие *дефицита* зоны  $Z^-(t)$ :  
 $D(t) = \sum_{v_i \in Z^-(t)} (rn - q_i(t))$ . То есть дефицит  $D(t)$  – это ресурс, которого зоне  $Z^-(t)$  не хватает до  $rn(n - k_t)$ .

**Утверждение 2.** При большом суммарном ресурсе  $W > T_2$  на любом такте  $t$  суммарные выходные (входные) потоки всех вершин из зоны  $Z^+(t)$  совпадают.

**Доказательство.** Вершина из зоны  $Z^+(t)$  отдает полную пропускную способность  $r$  по всем ребрам. Тогда суммарный выходной поток каждой из таких вершин  $v_i$  равен

$$f_i^{out}(t) = rn.$$

Суммарный входной поток одной вершины  $v_i$  из зоны  $Z^+(t)$  складывается из ресурса, отданного ей на такте  $t$  всеми вершинами из зоны  $Z^+$ , и всеми вершинами из зоны  $Z^-$ :

$$f_i^{in}(t+1) = rk_t + \sum_{j \in A_t} f_j^{out}(t).$$

То есть получаем не зависящие от  $i$  значения  $f_i^{out}(t)$  и  $f_i^{in}(t+1)$  для всех вершин  $v_i$  из зоны  $Z^+(t)$ .

Из доказательства следует, что ресурс в каждой вершине из зоны  $Z^+(t)$  либо сохраняется, либо уменьшается на одну и ту же величину ( $f_i^{in}(t+1) \leq f_i^{out}(t)$ ). При этом в начальный момент времени вершины были расположены в порядке убывания имеющегося в них ресурса, значит, переход в зону  $Z^-$  может осуществляться только начиная с наименее наполненной вершины – т.е. с  $v_{k_0}$ . Учитывая, что вершины, попавшие в зону  $Z^-$ , покинуть эту зону не могут, формулируем следующее следствие из утверждения 2.

**Следствие 1.** На каждом такте вершины из зоны  $Z^+$  теряют одинаковое количество ресурса, таким образом оставаясь упорядоченными по убыванию.

В случае когда все вершины находятся в зоне  $Z^+(t)$ , т.е.  $k_t = n$ , формулы из доказательства утверждения 2 переписы-



ваются следующим образом:

$$\forall i \in \overline{1, n} : f_i^{out}(t) = rn, f_i^{in}(t+1) = rn.$$

Получив, что суммарные входной и выходной потоки каждой вершины одинаковы и равны  $rn$ , формулируем второе следствие.

**Следствие 2.** *Если все вершины находятся в зоне  $Z^+(t)$ , то состояние  $Q(t)$  устойчиво.*

Пусть мощность зоны  $Z^+(t)$  постоянна, т.е.  $\forall t k_t = k$ . Такая ситуация имеет место при выполнении неравенства  $q_k(0) - rn > \frac{D(0)}{k}$ , иными словами, если «профицит» самой последней вершины  $v_k$  (вершины с наименьшим начальным ресурсом) из зоны  $Z^+(0)$  больше  $\frac{D(0)}{k}$ . Действительно, суммарный ресурс, который отдадут вершины из зоны  $Z^+(0)$ , не может превышать  $D(0)$ , так как зона  $Z^-(0)$  не может получить ресурса больше, чем  $D(0)$ .

**Утверждение 3.** *При большом суммарном ресурсе  $W > T_2$ , если выполняется неравенство  $q_k(0) - \frac{D(0)}{k} > rn$ , то предельное состояние сети имеет вид:*

$$Q^* = \left( q_1(0) - \frac{D(0)}{k}, \dots, q_k(0) - \frac{D(0)}{k}, rn, \dots, rn \right).$$

Доказательство см. в приложении.

Рассмотрим теперь общий случай:  $k_t$  непостоянно. Это, в свою очередь, эквивалентно тому, что вершины из зоны  $Z^+$ , начиная с наименее наполненной  $v_{k_0}$ , могут перейти в зону  $Z^-$ . Таким образом, мощность зоны  $Z^+$  невозрастает с течением времени.

Случай постоянной мощности рассмотрен в утверждении 3. Если же мощность уменьшается, то в некоторый момент времени  $t'$  этот процесс закончится, так как зона  $Z^+$  при  $W > T_2$  не может быть пуста. Тогда для любого  $t \geq t'$  имеем  $l > 0$  вершин в зоне  $Z^+(t)$  и  $n - l$  вершин в зоне  $Z^-(t)$ . То есть начиная с момента времени  $t'$  мощность зоны  $Z^+(t')$  является постоянной. Таким

образом, условия утверждения 3 для  $k = l$  и  $t'$  выполняются и сеть будет иметь следующее предельное состояние:

$$Q^* = \left( q_1(t') - \frac{D(t')}{l}, \dots, q_l(t') - \frac{D(t')}{l}, rn, \dots, rn \right).$$

Остается только найти состояния вершин зоны  $Z^+(t')$  ( $q_i(t')$ ,  $i = \overline{1, l}$ ) и дефицит  $D(t')$ .

Используя утверждение 2, заключаем, что все вершины из зоны  $Z^+(t)$  на каждом такте  $t$  отдают одинаковое количество ресурса зоне  $Z^-(t)$ , а значит, состояние каждой вершины  $v_j$ ,  $j = \overline{1, l}$ , к моменту времени  $t'$  суммарно уменьшилось на определенную величину  $\delta$ , т.е.

$$\forall j = \overline{1, l} : q_j(t') = q_j(0) - \delta.$$

По закону сохранения ресурса в сети, суммарное количество ресурса в вершинах  $v_i$ ,  $i = \overline{l+1, n}$ , к моменту времени  $t'$  увеличилось ровно на столько, сколько потеряли вершины  $v_j$ ,  $j = \overline{1, l}$ :

$$\sum_{i=l+1}^n q_i(t') = \sum_{i=l+1}^n q_i(0) + \delta l.$$

Тогда

$$D(t') = \sum_{i=l+1}^n (rn - q_i(t')) = \sum_{i=l+1}^n (rn - q_i(0)) - \delta l,$$

а предельное состояние вершин  $v_j$ ,  $j = \overline{1, l}$ , выглядит следующим образом:

$$\forall j = \overline{1, l} : q_j^* = q_j(t') - \frac{D(t')}{l} = q_j(0) - \frac{1}{l} \sum_{i=l+1}^n (rn - q_i(0)).$$

Итак,

$$\forall j = \overline{1, l} : q_j^* = q_j(0) - \frac{w}{l}, \quad w = \sum_{i=l+1}^n (rn - q_i(0)).$$

Здесь  $l$  – наибольшее целое число, для которого выполняется неравенство  $q_l^* > rn$ .

Таким образом, можно сформулировать следующую теорему.

**Теорема 1.** При большом суммарном ресурсе  $W > T_2$  предельное состояние сети существует и равно

$$Q^* = \left( q_1(0) - \frac{w}{l}, \dots, q_l(0) - \frac{w}{l}, rn, \dots, rn \right),$$

где  $w = \sum_{i=l+1}^n (rn - q_i(0))$ , а  $l$  – наибольшее целое число, для которого выполняется неравенство

$$q_l(0) - \frac{1}{l} \sum_{i=l+1}^n (rn - q_i(0)) > rn.$$

**Замечание 1.** Для случая постоянной мощности зоны  $Z^+$  (т.е.  $\forall t \ k_t = k$  и, следовательно,  $k^* = l = k$ ) имеем  $w = \sum_{i=k+1}^n (rn - q_i(0)) = D(0)$ . Поэтому утверждение 3 есть частный случай теоремы.

**Пример 1.** Рассмотрим полную однородную ресурсную сеть с пятью «жадными» вершинами и единичной пропускной способностью (рис. 2).

Начальное состояние подобрано таким образом, чтобы все вершины были уже насыщенными:

$$Q(0) = (14, 12, 7, 2, 1).$$

Суммарный ресурс  $W=36$  удовлетворяет условию  $W > rn^2=25$  зоны «достаточного большого» ресурса. Для нахождения предельного состояния сети воспользуемся полученной выше теоремой.

Найдем  $l$  – наибольшее целое число, для которого выполняется неравенство

$$q_l(0) - \frac{1}{l} \sum_{i=l+1}^n (rn - q_i(0)) > rn.$$

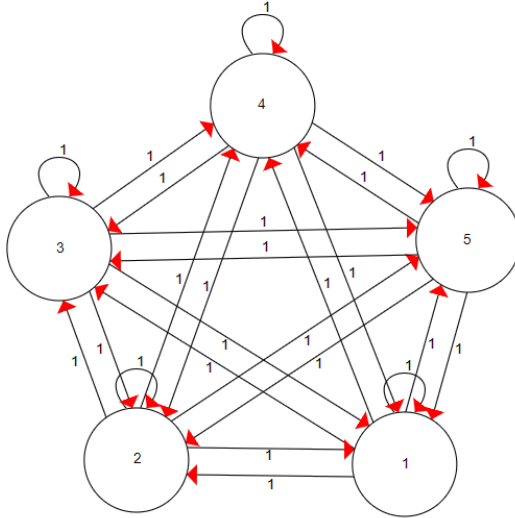


Рис. 2. Однородная ресурсная сеть с пятью «жадными» вершинами,  $r = 1$

$$l = 1 \Rightarrow$$

$$14 - (5 - 12 + 5 - 7 + 5 - 2 + 5 - 1) = 16 > 5 \text{ — неравенство выполняется.}$$

$$l = 2 \Rightarrow$$

$$12 - \frac{1}{2}(5 - 7 + 5 - 2 + 5 - 1) = \frac{19}{2} > 5 \text{ — неравенство выполняется.}$$

$$l = 3 \Rightarrow$$

$$7 - \frac{1}{3}(5 - 2 + 5 - 1) = \frac{14}{3} > 5 \text{ — неравенство не выполняется.}$$

Получили  $l = 2$ , т.е. в зоне  $Z^{+*}$  будут только вершины  $v_1, v_2$ .

Тогда предельное состояние сети вычисляется по формуле

$$Q^* = \left( q_1(0) - \frac{w}{2}, q_2(0) - \frac{w}{2}, rn, rn, rn \right).$$

Здесь  $w = \sum_{i=3}^5 (rn - q_i(0)) = 5 - 7 + 5 - 2 + 5 - 1 = 5$ .

Имеем следующее предельное состояние сети:

$$Q^* = (11, 5, 9, 5, 5, 5, 5).$$

Протокол работы сети приведен в таблице 1, динамика ресурса – на рис. 3. •

Таблица 1. Протокол работы сети с начальным состоянием  $Q(0) = (14, 12, 7, 2, 1)$

$t_i$	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$
0	14,0000	12,0000	7,0000	2,0000	1,0000
1	12,2500	10,2500	5,2500	4,0000	4,2500
2	11,8125	9,8125	4,8125	4,8125	4,7500
3	11,6563	9,6563	4,8906	4,8906	4,9063
4	11,5781	9,5781	4,9492	4,9492	4,9453
5	11,5391	9,5391	4,9736	4,9736	4,9746
6	11,5195	9,5195	4,9871	4,9871	4,9868
7	11,5098	9,5098	4,9935	4,9935	4,9935
8	11,5049	9,5049	4,9967	4,9967	4,9967
9	11,5024	9,5024	4,9984	4,9984	4,9984
10	11,5012	9,5012	4,9992	4,9992	4,9992
11	11,5006	9,5006	4,9996	4,9996	4,9996
12	11,5003	9,5003	4,9998	4,9998	4,9998
13	11,5002	9,5002	4,9999	4,9999	4,9999
14	11,5001	9,5001	4,9999	4,9999	4,9999
15	11,5000	9,5000	5,0000	5,0000	5,0000
16	11,5000	9,5000	5,0000	5,0000	5,0000
...	...	...	...	...	...

#### 4. Заключение

Рассмотрена модификация модели ресурсной сети – сеть с «жадными» вершинами. «Жадные» вершины на каждом такте сначала передают необходимое количество ресурса себе в петлю, а затем оставшийся ресурс (если он имеется) распределяют

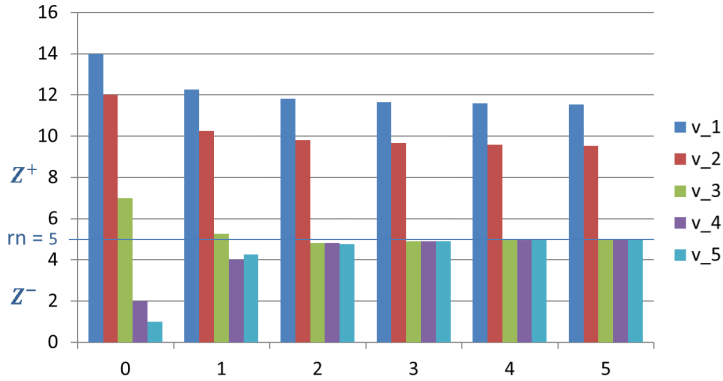


Рис. 3. Динамика ресурса в сети с начальным состоянием  $Q(0) = (14, 12, 7, 2, 1)$ .

в исходящие рёбра по правилам стандартной ресурсной сети. Исследование ограничено топологией полной однородной сети, т.е. сети, представленной полным графом, у которого веса ребер одинаковы.

Для такой модели известны пороговые значения  $T_1 = rn$  и  $T_2 = rn^2$ , разграничивающие зоны различного поведения сети: зоны «недостаточного» и «достаточного» ресурса, зоны «достаточного малого» и «достаточного большого» ресурса соответственно.

В данной работе описано функционирование сети, суммарный ресурс которой находится в зоне «достаточного большого» ресурса. Доказана сходимость процесса перераспределения ресурса, найдено предельное состояние сети.

В дальнейшем планируется исследование других топологий сети модели «жадных» вершин; нахождение основных характеристик функционирования, определение условий остановки сети, получение векторов предельных состояний и потоков для разных значений суммарного ресурса и различного начального распределения этого ресурса между вершинами.

## Литература

1. АБДУЛРАХМАН Х., ЕРУСАЛИМСКИЙ Я.М., СКОРОХОДОВ В.А. *Ресурсные сети с вентиляющей достижимостью* // Инженерный вестник Дона. – 2018. – №4. – С. 78.
2. АБДУЛРАХМАН Х., СКОРОХОДОВ В.А. *Ресурсные сети с магнитной достижимостью* // Известия ВУЗов. Северо-Кавказский регион. Естественные науки. – 2016. – №4. – С. 4–10.
3. АБДУЛРАХМАН Х., СКОРОХОДОВ В.А. *Полные двухресурсные сети с петлями* // Известия ВУЗов. Северо-Кавказский регион. Естественные науки. – 2016. – №2. – С. 10–16.
4. АБДУЛРАХМАН Х., СКОРОХОДОВ В.А. *Двухресурсные сети с магнитной достижимостью* // Известия ВУЗов. Северо-Кавказский регион. Естественные науки. – 2020. – №3. – С. 4–10.
5. АГАЕВ Р.П., ЧЕБОТАРЕВ П.Ю. *Сходимость и устойчивость в задачах согласования характеристик (обзор базовых результатов)* // Управление большими системами. – 2010. – №30.1 – С. 470–505.
6. АГАЕВ Р.П., ЧЕБОТАРЕВ П.Ю. *Метод проекции в задаче о консенсусе и регуляризованный предел степеней стохастической матрицы* // Автоматика и телемеханика. – 2011. – №12. – С. 38–59.
7. ГУБАНОВ Д.А., НОВИКОВ Д.А., ЧХАРТИШВИЛИ А.Г. *Социальные сети. Модели информационного влияния, управления и противоборства*. – М.: Изд-во физ.-мат. литры, 2010. – 228 с.
8. ЕРУСАЛИМСКИЙ Я.М., СКОРОХОДОВ В.А., КУЗЬМИНОВА М.В., ПЕТРОСЯН А.Г. *Графы с нестандартной достижимостью: задачи, приложения*. – Ростов-на-Дону: Южный федеральный университет, 2009. – 195 с.
9. ЖИГИРЬ А.А. *Методы количественной оценки экономического риска строительной организации при реализации ин-*

- вестиционных проектов // ЕГИ. – 2020. – №1. – С. 106–111.
10. ЖИЛЯКОВА Л.Ю., КУЗНЕЦОВ О.П. *Теория ресурсных сетей: монография.* – М.: РИОР: ИНФРА-М, 2017. – 283 с.
  11. ЖИЛЯКОВА Л.Ю., ЧАПЛИНСКАЯ Н.В. *Исследование полных однородных ресурсных сетей с «жадными» вершинами* // Управление большими системами (в печати).
  12. ЖИЛЯКОВА Л.Ю. *Ресурсные сети с ограничениями на ёмкость вершин* [Электронный ресурс]: монография. – М.: ИЦ РИОР, НИЦ ИНФРА-М, 2018. – 160 с.
  13. КУЗНЕЦОВ О.П. *Однородные ресурсные сети. I. Полные графы* // Автоматика и телемеханика. – 2009. – №11. – С. 136–147.
  14. СКОРОХОДОВ В.А., АБДУЛРАХМАН Х. *Динамические ресурсные сети. Случай малого ресурса* // Вестник Воронежского государственного университета. Серия: Физика. Математика. – 2018. – №4. – С. 186–194.
  15. BLANCHARD PH., VOLCHENKOV D. *Random Walks and Diffusions on Graphs and Databases: An Introduction (Springer Series in Synergetics).* – Berlin – Heidelberg: Springer-Verlag, 2011. – 262Berlin – Heidelbergp.
  16. MASUDA N., PORTER M.A., LAMBIOTTE R. *Random walks and diffusion on networks* // Physics Reports. – 2017. – Vol. 716–717. – P. 1–58.

## Приложение

Доказательство утверждения 3.

«Профицит» вершины  $v_k$  из зоны  $Z^+(0)$  превосходит величину  $\frac{D(0)}{k}$ , тогда в любой момент времени  $t$  в сети имеется ровно  $k$  вершин из зоны  $Z^+$ .

Рассмотрим разность дефицитов на тактах  $t + 1$  и  $t$ :

$$D(t + 1) - D(t) = \sum_{i=k+1}^n (rn - q_i(t + 1)) - \sum_{i=k+1}^n (rn - q_i(t)) =$$



$$\begin{aligned}
 &= \sum_{i=k+1}^n q_i(t) - \sum_{i=k+1}^n q_i(t+1) = Q^-(t) - Q^-(t+1) = \\
 &= Q^-(t) - \left( Q^-(t) - \sum_{i=k+1}^n \frac{k(q_i(t) - r)}{n-1} + kr(n-k) \right) = \\
 &= \sum_{i=k+1}^n \frac{k(q_i(t) - r)}{n-1} - kr(n-k) = \frac{k}{n-1} Q^-(t) - \frac{krn(n-k)}{n-1}.
 \end{aligned}$$

С другой стороны,

$$D(t) = \sum_{i=k+1}^n (rn - q_i(t)) = rn(n-k) - Q^-(t),$$

$$Q^-(t) = rn(n-k) - D(t).$$

Тогда имеем

$$D(t+1) - D(t) = \frac{k}{n-1} (rn(n-k) - D(t)) - \frac{krn(n-k)}{n-1},$$

$$D(t+1) = \frac{n-k-1}{n-1} D(t),$$

$$D(t) = \left( \frac{n-k-1}{n-1} \right)^t D(0).$$

Получили убывающую геометрическую прогрессию, а, значит,

$$\lim_{t \rightarrow \infty} D(t) = 0.$$

Таким образом, зона  $Z^{-*}$  не будет иметь дефицита, а, учитывая тот факт, что вершины не могут покидать зону  $Z^-(t)$ , заключаем, что в предельном состоянии все вершины  $v_i$ ,  $i = k+1, n$ , имеют  $rn$  ресурса.

Теперь найдем предельный ресурс в первых  $k$  вершинах – вершинах зоны  $Z^+(t)$ . По утверждению 2, выходные и входные потоки вершин зоны  $Z^+(t)$  будут совпадать, значит, ресурс у этих вершин меняется одинаково. Суммарно отданный зоне  $Z^-(t)$  ресурс должен быть равен дефициту. Тогда получим, что ресурс в каждой вершине зоны  $Z^+(t)$  уменьшился на  $\frac{1}{k} D(0)$ .

Итого, предельное состояние сети выглядит следующим образом:

$$Q^* = \left( q_1(0) - \frac{D(0)}{k}, \dots, q_k(0) - \frac{D(0)}{k}, rn, \dots, rn \right).$$

## **RESEARCH OF COMPLETE HOMOGENEOUS "GREEDY-VERTICES" RESOURCE NETWORKS: ZONE OF "SUFFICIENT LARGE" RESOURCE**

**Nadezda Chaplinskaya**, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, student (nadya1462@gmail.com).

*Abstract: "Resource network with greedy vertices" – the modification of the graph dynamic model "resource network". At each discrete-time moment the graph vertices transfer resources to each other through the edges with limited throughput, first passing the available resource to themselves via the loop and then distributing the remaining resource (if it is available) to outgoing edges according to the "standard" resource network rules. These are two rules with threshold switching: if the vertex has resource, that exceeds the total throughput of all vertex's outgoing edges, it transfers the full throughput to each outgoing edge; otherwise, it gives away the entire available resource, distributing it in proportion to the throughputs of the outgoing edges. The particular case of a complete homogeneous resource network with "greedy" vertices is considered. For networks of such type there are two total resource thresholds, separating zones of different network behaviour: the first threshold divides the zones of "insufficient" and "sufficient" resources, the second divides the zones of "sufficient small" and "sufficient large" resources. In the article the last zone – "sufficient large" resource – is investigated: the functioning of the network is described, the asymptotic state is found.*

Keywords: graph dynamic threshold model, resource net, "greedy-vertices" model, zone of "sufficient large" resource.

УДК 519.1

ББК 22.176

DOI: 10.25728/ubs.2021.90.3

*Статья представлена к публикации  
членом редакционной коллегии О.П. Кузнецовым.*

*Поступила в редакцию 29.01.2021.*

*Дата опубликования 31.03.2021.*

## К ВОПРОСУ О КОРРЕКТИРОВКЕ ПОСТАНОВКИ УПРАВЛЕНЧЕСКОЙ ЗАДАЧИ

Шевыренков М. Ю.<sup>1</sup>

(ФГБУН Институт проблем управления  
им. В.А. Трапезникова РАН, Москва)

*Рассматривается проблема уменьшения влияния негативных факторов на корректность постановки задачи. Поскольку эти факторы обладают сложной, комплексной природой, в работе представлен подход, основанный на использовании интеррогативной логики, который, по мнению автора, способен уменьшить степень влияния рассматриваемых факторов на эффективность принятия управленческих решений. В рамках создания такого подхода решён ряд задач. Для решения проблем, связанных с корректностью постановки управленческих задач при активном прогнозировании, разработан механизм корректировки постановки управленческих задач для информационного управления социально-экономическими системами, основанный на аппарате интеррогативной логики. Сформулирован и обоснован список необходимых условий для корректной постановки управленческой задачи при информационном управлении социально-экономическими системами при помощи активных прогнозов. Также сформулирован специальный формально-содержательный интеррогативный язык, использующийся для уменьшения потерь информации при передаче задания исполнителю.*

Ключевые слова: интеррогативная логика, управленческие задачи, корректность постановки задачи, стратегический менеджмент, стратегическое планирование, принятие решений, управленческая задача.

### 1. Введение

С начала 1980-х годов крупные предприятия по всему миру ведут свою производственную деятельность в постоянно меняющихся условиях. Изменение структуры национальных экономик, появление новых технологий, а также постоянно меняющийся характер потребительских предпочтений приводят к тому, что управляющим предприятиями нужно постоянно адаптироваться к этим переменам: предвидеть их и принимать соответствующие управленческие решения [9]. Это особенно хорошо видно на примере предприятий, чьи продукты имеют длительные периоды окупаемости и требуют значительных инвестици-

---

<sup>1</sup> Максим Юрьевич Шевыренков, аспирант (sheverenkov@mail.ru).

онных ресурсов, в частности в аэрокосмической и нефтегазовых отраслях. В этих отраслях цена ошибки при постановке задачи планирования или управления предприятием особенно высока. Поэтому постановка управленческой задачи уже давно признана важной компонентой принятия стратегических решений [22, 28]. В данной работе для управленческой задачи будет использоваться определение из работы Баэра, Диркса и Никельсона [10] «Мы определяем (стратегическую) задачу как отклонение от желаемого набора конкретных или ряда приемлемых условий, приводящее к появлению одной или нескольких связанных проблем, которые необходимо решить<sup>1</sup>».

Причинами такого отклонения могут служить различные факторы, приведём некоторые из них:

1. Вызванные определённым видом неопределённости, в которой находится управляющий субъект.

2. Вызванные тем, что управленческое решение связано с некоторым известным или неизвестным распределением рисков между различными вариантами принятия данного решения.

3. Вызванные особенностями мышления управляющих субъектов (когнитивные искажения, культурные особенности).

4. Вызванные отсутствием корректной постановки задачи или формулировки цели, заданной управляющим субъектом.

В современной научной литературе необходимо отметить работы Айзенхардта [11], Баэра, Диркса и Никельсона [10] и Маритана [20] которые изучают различные аспекты постановки и решения стратегических управленческих задач. К сожалению, в настоящий момент всё ещё продолжается дискуссия о том, какую «метрику» следует использовать для оценки корректности поставленной задачи. Баэр, Диркс и Никельсон предлагают использовать концепт «полноты» (*comprehensiveness*) для оценки того, насколько одна постановка задачи лучше или хуже другой. По сути, они предлагают сравнивать количество учтённых

---

<sup>1</sup> В опус.: «We define a (strategic) problem as a deviation from a desired set of specific or a range of acceptable conditions resulting in a symptom or a web of symptoms recognized as needing to be addressed».

при постановке задачи проблем. Данный подход представляется автору перспективным, но требующим доработки.

В настоящей работе автор развивает вышеупомянутый подход Баэра, Диркса и Никельсона с применением формального аппарата интеррогативной логики. Рассмотрим этот вид логики более подробно.

В контексте рассматриваемой статьи интеррогативная логика будет интересовать нас в первую очередь как инструмент формирования формальных имитаций вопросов естественного языка для уточнения имеющихся вопросов.

В процессе решения экспертных задач третьего и четвертого уровня сложности [6] в области управления часто возникает ряд трудностей, препятствующих их эффективному решению. Они могут быть связаны как с особенностями мышления человека, например когнитивными искажениями<sup>1</sup>, так и с особенностями решаемых управленческих задач, связанных с чрезвычайно большим количеством возможных вариантов их решения, управлением несколькими различными по своему составу, строению и мотивации рабочими коллективами. Например, одной из таких особенностей является необходимость корректно сформулировать постановку общей задачи для разных коллективных субъектов, работающих в рамках одной организации, но имеющих различную мотивацию и цели, которые зачастую противостоят друг другу [18].

Также необходимо упомянуть столь тривиальную, но крайне распространенную ошибку, затрудняющую управление предприятиями, как неверная оценка объемов работ, необходимых для реализации проекта и времени на их выполнение. Особенно хорошо она видна на примере фирм, работающих на молодых, развивающихся рынках или участвующих в создании уникальных крупных инфраструктурных или технически сложных объектов [26]. В таких случаях первоначальные формулировки постановки задачи часто содержат вышеуказанные ошиб-

---

<sup>1</sup> *Которые, будучи ошибками в суждениях, могут привести к принятию неверного решения даже в случае если управленец обладает полными и достоверными данными об имеющейся ситуации [1].*

ки, которые можно устранить в процессе корректировки формулировок постановки этих задач.

С целью уменьшения степени влияния вышеуказанных проблем на качество решений управленческих задач представлен подход к созданию нового метода повышения степени корректности постановки управленческих задач на основе интеррогативной логики. Представленный подход, по мнению автора, позволит повысить уровень эффективности решения ряда управленческих задач.

Целью данной работы является формулировка и доказательство необходимых условий для корректной постановки управленческой задачи при прогнозировании и планировании развития социально-экономических систем на основе интеррогативной логики.

Для реализации поставленной исследовательской цели декомпозируем ее и рассмотрим вспомогательные следующие задачи:

1. Ввести и обосновать новый язык исчисления предикатов второго порядка  $\text{Яи}_1$ , на основе которого возможно уменьшить влияние указанных ниже проблем на постановку управленческих задач.

2. Ввести на основе нового языка исчисления предикатов второго порядка  $\text{Яи}_1$  строго формальное определение некорректных вопросов, которое бы отвечало дополнительным условиям по сравнению с аналогичным введенным Е.К. Войшвилло и Ю.А. Петровым [5] и расширяло бы применение языка  $\text{Яи}$  для решения практических задач.

## ***2. Подход к созданию корректной постановки задачи***

В данном случае автору настоящей работы представляется перспективным использовать для корректировки постановки задач в виде вопросов формализованный подход с использованием эротетической (интеррогативной) логики, по сути, логики вопросов и ответов. В контексте этой работы интеррогативная логика будет интересовать нас в первую очередь как инструмент формирования формальных имитаций вопросов естественного языка для уточнения имеющихся вопросов. Это позволит более

корректно ставить управленческие задачи при создании активных прогнозов.

Логика вопросов и ответов зародилась во второй половине семидесятых годов XIX века. Одним из её основоположников является британский логик Ричард Уэйтли [24]. Ему принадлежит заслуга того, что он одним из первых применил аппарат формальной логики для анализа вопроса. В тридцатые годы прошлого века в развитии эротетической логики принимал участие также Евгений Сперэнция [23].

Послевоенный период и время появления первых ЭВМ ознаменовались значительным усилением внимания к логике вопросов и ответов, к этому периоду относятся работа Артура и Мэри Прайор в 1955 г. [21], Чарльза Хемблина [13]. Другой заметной работой, специально посвященной логике вопросов и ответов, была работа Тадеуша Кубиньского [19]. Говоря о более поздних вкладах в логику вопросов и ответов, необходимо упомянуть работы Дэвида Харра [14], а также финского учёного Яакко Хинтикка (1962 г.) [16], применившего эпистемическую логику для исследования вопросов, и монографию Нормана Белнапа и Тимоти Стила [2]. Среди современных работ можно отметить исследование Анджея Вишневого [25], Яцина Хамама [12].

В российской литературе можно отметить работы Владимира Федотовича Беркова [3], Евгения Казимировича Войшвилло и Юрия Александровича Петрова [5], а также Бориса Ивановича Федорова [7].

Разработанная Хинтиккой модель интеррогативного запроса (IMI – interrogative model of inquiry) [15, 17] представляет несомненный интерес для изучения процесса корректировки постановки задачи. Так, IMI позволяет описать процесс корректировки некорректной формулировки постановки задачи, когда Эксперт путём логических выводов и опроса ЛППР пытается выяснить, что именно нужно его заказчику ЛППР. Такой способ решения проблемы некорректной формулировки постановки задачи в неформализованном виде часто применяется на практике, но, по мнению автора, не является наилучшим. Хорошим примером последствий такого подхода является процесс утвер-

ждения проекта британского штурмового танка во время второй мировой войны, который из-за плохо поставленной задачи затянулся на 10 месяцев, за время которых было разработано 18 вариантов проекта, в результате чего танк был построен уже после войны [27]. Попытки подрядчиков, выступавших в роли Экспертов, выяснить желания заказчика привели к растрате средств на разработку подобного танка и к тому, что на завершающем этапе войны при прорыве немецких укреплений британская армия осталась без современных танков прорыва, что заметно ограничило её тактические возможности.

В данной же работе автор пытается решить проблему корректировки постановки задачи путём формирования набора требований к формулировке постановки задачи, которые позволят минимизировать количество итераций в процессе уточнения формулировки постановки задачи.

Из рассмотренных публикаций по логике вопросов и ответов наибольший интерес в контексте корректировки постановки задач для нас представляет работа Е.К. Войшвилло и Ю.А. Петрова [5], поскольку в ней вводится и рассматривается контекстуальное определение некорректного вопроса, что чрезвычайно важно для корректировки постановки задач в виде вопросов, например: «Как нам достигнуть заданных темпов роста?» или «Как с помощью активного прогноза повлиять на экономический рост?».

В этой же работе Е.К. Войшвилло и Ю.А. Петрова авторами предложен язык исчисления предикатов второго порядка Яи, который можно использовать для формализации формулировок постановок задачи в виде вопроса. В рамках такой формализации, т.е. в случае перевода формулировки постановки задачи с естественного языка на формальный язык Яи, отсеивается большое количество тривиально-некорректных или неосмысленных вопросов, на которые невозможно дать ответ, не задавая уточняющих вопросов, например: «Стол?», «Повысить ли белый?», «Изменить ли ставку реформирования налоговой нагрузки?» и т.д.

В лучшем варианте для ЛПП за счёт этого экономится значительное количество времени, которое в противном случае было бы потрачено на уточнение вопроса.



Применение языка Яи позволяет решить ряд проблем, связанных с корректировкой постановок задач в виде вопроса, однако остаётся достаточно широкий круг задач, которые возникают при корректировке постановки задач и которые затруднительно решить при помощи использования языка, введённого Войшвилло и Петровым.

В частности, эти затруднения возникают из-за того, что его авторы, во-первых, рассматривают идеальный случай без учёта того фактора, что на вопрос могут отвечать разные агенты, для некоторых из них формулировка вопроса может быть корректной, а для других идентичная ей формулировка – некорректной.

Во-вторых, авторы дают лишь описательное определение некорректных вопросов: Причины тривиальной некорректности в естественном языке могут состоять в неправильном построении самого вопросительного предложения или в неосмысленности каких-либо его частей, например, употребляемых в нём предикатов. Предикат может не иметь определяемого смысла, если он не определён или если какой-либо предикат, хотя и определённый на некоторой области предметов, применяется к предметам из другой области. Так, неосмысленным и, значит, тривиально некорректным можно считать вопрос: «Существуют ли чётные трансфинитные числа?», поскольку понятие чётности не определено на множестве трансфинитных чисел [5]. Мы же в дальнейшем введём строго формальное определение некорректных вопросов с опорой на материалы работы [2] и с учётом того, что на вопрос будут отвечать агенты с различным жизненным опытом, компетенциями и квалификацией. В практике управления возникает также ряд проблем, связанных с тем, что множество интерпретаций формулировки постановки задачи агента, ставящего задачу в виде вопроса, может не совпадать с множеством интерпретаций этой же формулировки агентом, который должен выполнять эту задачу. Например, на вопрос: «Каков целевой размер  $IRR^1$ ?» существует несколько вариантов ответа в зависимости от расшифровки аббревиатуры и соответ-

---

<sup>1</sup> Обозначение  $IRR$  в экономике имеет несколько значений: *implied repo rate* (предполагаемая ставка РЕПО), *internal rate of return* (внутренняя норма прибыли) и т.д.

ствующего контекста, а отвечающий агент может об этом не знать, например, в силу недостаточной квалификации.

### **3. Введение нового языка исчисления предикатов второго порядка Яи<sub>1</sub>**

Для уменьшения влияния указанных проблем на постановку управленческих задач в настоящем исследовании вводится язык Яи<sub>1</sub>.

Перед началом описания языка Яи<sub>1</sub> дадим пояснение понятий вопроса и ответа. Учитывая специфику решаемых задач, автору представляется целесообразным использовать «компьютерный» подход к определению понятия вопрос, близкий к используемому у Белнапа и Стила [2]. Согласно этому подходу, вопрос является инструкцией для отвечающего, который должен устранить указанную в вопросе неполноту знаний спрашивающего. Так, хорошим примером может послужить управленческая задача в энергетике в виде вопроса: «Можно ли полностью обеспечить возобновляемой энергией все потребности населения страны?».

В данной работе вопрос рассматривается как высказывание, в котором существует неопределённость относительно истинности одной или нескольких содержащихся в нём переменных и/или отношений между ними. Для обозначения неопределённости значения истинности переменной/отношения между переменными вводится вопросительный оператор «?».

Под ответом на вопрос понимается высказывание относительно истинности или ложности переменной или отношения, обозначенного вопросным оператором.

Язык Яи<sub>1</sub> является расширением содержательного интеррогативного языка Яи, представленного в работе [5]. В языке Яи<sub>1</sub> будет пересмотрена и изменена трактовка некорректных вопросов, рассмотренных в языке Яи, а также будет добавлен ряд новых правил, призванных адаптировать язык Яи для корректировки формулировок постановки практических управленческих задач.

В начале представим часть описания элементов языка Яи, являющихся общими для языков Яи и Яи<sub>1</sub>.

Исходя из стандартной схемы построения формализованного языка сначала зададим его алфавит (см. например [4, с. 39]). В него включены:

1. Логические символы – знаки для логических терминов. Символ импликации  $\supset$ , вопросительного оператора  $?$ , отрицания  $\neg$ , конъюнкции  $\&$ , дизъюнкции  $\vee$ , кванторы всеобщности  $\forall$  и существования  $\exists$ .

2. Нелогические символы предназначены для замещения простых высказываний или нелогичных терминов различных категорий. К ним в этом языке относятся следующие:  $x_1, x_2, x_3, \dots$  – предметные (индивидуальные) переменные;  $a_1, a_2, a_3, \dots$  – предметные (индивидуальные) постоянные;  $P_n, P_{n1}, P_{n2}, \dots$  – предикатные постоянные (предикаты);  $X_n, X_{n1}, X_{n2}, \dots$  – предикатные переменные, где  $n$  – число мест предиката,  $n = 1, 2, \dots$ . Как обычно, предметные переменные и предметные постоянные назовем терминами.

3. Технические символы, например скобки.

Теперь рассмотрим правила синтаксиса, под которыми мы понимаем правила образования выражений данного языка из исходных символов. В таком случае под выражениями языка понимаются формулы, посредством которых фиксируется логические формы высказываний.

Образование формул в языке Яи определяется синтаксическими правилами образования<sup>1</sup>, которые разбиваются на две группы: согласно одним (ПЯ<sub>и</sub>1 – ПЯ<sub>и</sub>3) образуются формулы, называемые пропозициональными, согласно другим (ПЯ<sub>и</sub>4 – ПЯ<sub>и</sub>6) – вопросительные формулы (вопросы).

ПЯ<sub>и</sub>1. Если  $\Delta^n$  –  $n$ -местный предикатный символ (предикатная переменная или постоянная) и  $t_1, t_2, \dots, t_n$  – термы (не обязательно различные), то  $\Delta^n(t_1, \dots, t_n)$  – пропозициональная формула<sup>2</sup>.

ПЯ<sub>и</sub>2. Если  $A, B$  — пропозициональные формулы, то  $\neg A, (A \& B), (A \vee B), (A \supset B)$  – пропозициональные формулы.

---

<sup>1</sup> ПЯ<sub>и</sub> – правила языка интеррогативного с числовыми индексами, обозначающими номер правила.

<sup>2</sup> В дальнейшем для краткости записи мы не будем указывать число мест предикатной переменной или постоянной.

ПЯ<sub>н</sub>3. Если  $A$  – пропозициональная формула и  $\gamma$  – предметная или предикатная переменная, то  $\forall\gamma A$ ,  $\exists\gamma A$  – пропозициональные формулы.

Свободной переменной мы считаем такую переменную, входящую в формулу, что она не следует непосредственно за квантором и не входит в область действия квантора по этой переменной. Все прочие переменные, входящие в формулу, будем называть связанными.

ПЯ<sub>н</sub>4. Если  $A$  – пропозициональная формула и  $\gamma_1, \dots, \gamma_m$  – все имеющиеся в ней различные свободные переменные, то  $(\gamma_1, \dots, \gamma_m)A$  есть вопрос. (При  $m = 1$  скобки, в которые заключены  $\gamma_1, \dots, \gamma_m$ , опускаются.) Вопрос типа: «Для каких  $\gamma$  истинно  $A$ ?» Например, высказывание  $A$  вида: «Спортсмен, занявший призовое место, но не ставший победителем, получил поощрительный приз.» будет иметь следующую формульную запись в виде простых высказываний:  $(\neg\gamma_1 \ \& \ \gamma_2) \supset \gamma_3$ , где переменные обозначают следующие простые высказывания:  $\gamma_1$  – спортсмен стал победителем соревнований,  $\gamma_2$  – спортсмен стал призёром соревнований,  $\gamma_3$  – спортсмен получил поощрительный приз. В этом случае на вопрос вида  $(\gamma_1, \gamma_2, \gamma_3)A$  можно дать ответ, что  $A$  истинно для  $\gamma_2$  и  $\gamma_3$ , но ложно для  $\gamma_1$ .

ПЯ<sub>н</sub>5. Если  $A$  – пропозициональная формула, содержащая одну свободную переменную  $\gamma$ , то  $(\forall\gamma)A$  и  $(\exists\gamma)A$  – вопросы. Вопросы типа: для всех ли  $\gamma$  истинно  $A$  и существует ли  $\gamma$ , для которого истинно  $A$ .

ПЯ<sub>н</sub>6. Если  $A$  – пропозициональная формула, не содержащая свободных переменных, и  $\varepsilon$  есть входящая в нее предметная или предикатная постоянная, то  $(?\varepsilon)A$  – вопрос. Вопрос типа: верно ли, что для  $\varepsilon$  истинно  $A$ ?

Пропозициональные формулы, не содержащие символов  $\neg$ ,  $\&$ ,  $\vee$ ,  $\supset$  будут называться простыми, остальные – сложными.

Правила ПЯ<sub>н</sub>4–6 определяют понятие простого вопроса, хотя бы и образованного из сложной пропозициональной формулы. Этого вида вопросы можно подразделить на прямые и непрямые вопросы. Прямые вопросы – вопросы, образованные по правилам ПЯ<sub>н</sub>5–6. Они имеют ещё название «ли-вопросов». Не-

прямые вопросы (или вопросы «для каких») – это вопросы, образованные по правилу ПЯ<sub>и</sub>4.

ПЯ<sub>и</sub>7. Если  $\alpha$  – пропозициональная формула,  $A, B$  – вопросы, то  $(\alpha \ \& \ A)$ ,  $(A \ \& \ B)$ ,  $(\alpha \ \vee \ A)$ ,  $(A \ \vee \ B)$ ,  $(\alpha \ \supset \ A)$  – вопросы, называемые (соответственно логическим связкам) конъюнктивным, дизъюнктивным и имплицативным вопросами. Все эти вопросы являются сложными. Рассмотрим это на примерах, где переменной  $\alpha$  соответствует высказывание «Сократ – человек», вопросу  $A$  – «Смертны ли люди?», вопросу  $B$  – «Собака ли Сократ?». Тогда вопрос  $(\alpha \ \& \ A)$  будет звучать как: «Верно ли, что Сократ – человек и люди смертны?», вопрос  $(A \ \& \ B)$  – как «Верно ли, что все люди смертны и Сократ – собака?», вопрос  $(\alpha \ \vee \ A)$  – как «Верно ли, что или Сократ – человек, или все люди смертны?», вопрос  $(A \ \vee \ B)$  – как «Верно ли, что или все люди смертны или Сократ – собака?», а вопрос  $(\alpha \ \supset \ A)$  – «Верно ли, что из-за того, что Сократ – человек, все люди смертны?».

Полагаем что, синтаксис языка Яи определен правилами ПЯ<sub>и</sub>1–7.

Теперь рассмотрим правила семантики языка Яи. Напомним, что под семантикой понимают «исследование отношений между знаками и обозначаемыми предметами, т.е. исследование значений и смыслов знаков» [5].

В качестве основы для интерпретации тех или иных выражений выбирается некоторое непустое множество индивидов (предметов)  $M$  и множество определенных на  $M$  отношений  $R$  (с различным числом мест). Кортеж  $S = \langle M, R \rangle$  назовем структурой. Интерпретация (на основе выбранной структуры  $\langle M, R \rangle$ ) дается с помощью следующих правил, согласно которым постоянным приписываются определенные значения, а переменным – лишь области возможных значений.

И1. Каждой (интерпретируемой) предметной постоянной сопоставляется в качестве ее значения предмет из  $M$ .

И2. Каждой  $n$ -местной предикатной постоянной сопоставляется  $n$ -местное отношение из  $R$ .

И3. Областью значений предметных переменных является множество  $M$ .

И4. Предикатные переменные принимают значения из  $R$  (причем значениями  $n$ -местной переменной могут быть лишь  $n$ -местные отношения).

В результате приписываний (интерпретаций по правилам И1–И3) все пропозициональные формулы, не содержащие свободных переменных, превращаются в предложения естественно-го языка. Все остальные пропозициональные формулы с учетом применения также И4 представляют собой предикатные выражения (называемые часто просто предикатами, или пропозициональными формулами), которые могут выполняться (принимать значение «истинно») при одних значениях переменных и не выполняться (принимать значение «ложно») при других.

В нижеследующей интерпретации формул-вопросов предполагается наличие какой-либо интерпретации пропозициональных формул в некоторой структуре  $S$ . Смысл простым вопросам будет придан следующими правилами И<sub>в</sub>С (интерпретация вопросов смысловая):

И<sub>в</sub>С1. Формула-вопрос  $(?(\gamma_1, \dots, \gamma_n))A$ , образованная по правилу ПЯ<sub>н</sub>4, имеет смысл, выражаемый предложением «Какие объекты из  $S$  удовлетворяют условию  $A$ ?», т.е. «Какие значения переменных  $\gamma_1, \dots, \gamma_n$  выполняют  $A$ ?».

И<sub>в</sub>С2. Формула-вопрос  $(?\chi \gamma) \chi A(\gamma)$ , построенная по правилу ПЯ<sub>н</sub>5, имеет смысл, передаваемый предложением: «Для всяких ли (или «существует ли» – в зависимости от того, есть ли  $\chi$  квантор  $\forall$  или  $\exists$ )  $\gamma$  такое, что выполняется  $A$ ?».

И<sub>в</sub>С3. Формула-вопрос  $(?\varepsilon)A(\varepsilon)$ , построенная по правилу ПЯ<sub>н</sub>6, имеет смысл, выражаемый предложением: «Верно ли  $A$  для объекта  $\varepsilon$ ?».

Отметим что синтаксис языка (ПЯ<sub>н</sub>1–7), правила интерпретации (И1–И7) и правила смысловой интерпретации И<sub>в</sub>С1–И<sub>в</sub>С3 языка Я<sub>и</sub> справедливы и для языка Я<sub>и</sub>1. Обозначения, применяемые в Я<sub>и</sub>1, остаются идентичными по форме и смыслу обозначениям, применяемым в языке Я<sub>и</sub>, если не указано обратного.

Как уже говорилось выше, основные изменения, отличающие Я<sub>и</sub>1 от Я<sub>и</sub>, лежат в синтаксической составляющей и связаны с изменённым определением понятия корректности вопроса.

В данной работе предлагается другая трактовка понятия корректности вопроса, которая, по мнению автора, расширяет применение языка Яи для решения практических задач.

Для этого введём новое понятие<sup>1</sup> «мир  $W$ », отсутствовавшее в языке Яи, и дадим формальное определение некорректных вопросов, которые в совокупности, по мнению автора, поможет учесть специфику практического применения языка Яи<sub>1</sub>.

В рамках рассматриваемого исследования под миром  $W$  будем понимать совокупность объектов или предметов<sup>2</sup> –  $M$ , отношений между ними –  $R$  и языка, который позволяет описывать этот мир, – Яи. При этом  $M \subset W$ ,  $R \subset W$  и  $S \subset W$ , множества  $M$  и  $R$  не пустые, где символ  $S$  обозначает кортеж  $\langle M, R \rangle$ .

Также рассмотрим два множества:  $Q$  – множество всех возможных вопросов, которые можно сформулировать существующими в мире  $W$  средствами этого языка, и  $O$  – множество всех возможных ответов на них, при этом верно, что пересечение  $O \cap W$  не пусто.

В частном случае такой мир может описывать производственную деятельность предприятия и его окружающую среду.

#### **4. Введение формального определения некорректных вопросов**

На основе нового языка исчисления предикатов второго порядка Яи<sub>1</sub> введём формальное определение некорректных вопросов, которое бы отвечало дополнительным условиям по сравнению с аналогичным введенным Е.К. Войшвилло и Ю.А. Петровым [11]. Необходимо оговориться, что в данном случае мы рассматриваем понятие корректности вопроса только для языка Яи<sub>1</sub>.

Перед тем как дать новое определение некорректности вопроса в рамках языка Яи<sub>1</sub>, поясним вышесказанное на поясняющем примере.

Одной из наиболее древних некорректных задач в виде вопроса ставилась правителями различных государств перед ал-

---

<sup>1</sup> В некотором роде аналогичное понятию «мир» приведенному в [7].

<sup>2</sup> При этом одним объектом является ЛПП и хотя бы одним – Эксперт.

химиками: «Можно ли превратить свинец в золото?». Учитывая то, что стоявшая перед правителями цель в виде увеличения количества денег в казне достигалась и достигается совершенно другими средствами<sup>1</sup>, несоответствие формулировки задачи имеющейся цели вполне очевидно и уже само по себе позволяет говорить о некорректности формулировки задачи.

Предположим, что существовало некое средневековое государство, которое, по сути, являлось своим миром, т.е. совокупностью предметов и отношений между ними, в том числе знаний, населяющих этот мир людей. Обозначим мир этого государства через  $W$ .

Поскольку процесс превращения одних химических элементов в другие при помощи нейтронной бомбардировки был открыт только в начале XX века, в средневековом мире  $W$  на вопрос о возможности трансмутации свинца в золото нельзя ответить ни утвердительно, ни отрицательно.

Поэтому, зная истинный ответ на рассматриваемый вопрос и то, что путь его получения был не связан с задачей по трансмутации золота, можно с определённой степенью утверждать, что корректность вопроса, зависит от того, принадлежит ли истинный ответ миру  $W$  или нет.

Тогда некорректными вопросами первого рода (или тривиально некорректными) будем называть такие элементы множества  $Q_1 \subseteq Q$ ,  $q_1 \in Q_1$ , являющегося подмножеством множества вопросов  $Q$ , множество истинных ответов на которые  $O_1 \notin W$ . При этом  $O_1 \subseteq O$ . Другими словами, такой вопрос, истинный ответ на который нельзя сформулировать на языке  $Я_1$  в мире  $W$ .

Если вернуться к нашему примеру, то вопрос правителя является некорректным вопросом первого рода, потому что истинный ответ был получен и сформулирован вне мира  $W$ .

Некорректными вопросами второго рода (нетривиально некорректные вопросы) будем называть такие элементы  $q_2$  множе-

---

<sup>1</sup> А именно, либо путём снижения содержания драгметаллов в монете или путём создания других видов денег, например бумажных, что и было сделано в Китае в X веке во времена империи Сун.



ства  $Q$ , множество ответов на которые  $O_2$  удовлетворяют двум следующим условиям:

1.  $O_2 \in W$ .

2. Элементы множества  $O_2$  как логические высказывания содержат хотя бы одну предикатную и/или предметную постоянную, которая не принадлежит множествам  $R$  или  $M$ , и/или хотя бы одну предикатную и/или предметную переменную, область значения которой не полностью принадлежит соответственно множествам  $R$  или  $M$ .

При этом  $O_2 \subseteq O$ .

Поясним возможность существования второго условия данного положения на следующем примере с золотом.

Современная наука позволяет достоверно ответить на вопрос о превращении металлов в золото, т.е. ответ на этот вопрос уже находится в мире  $W$  в виде необходимых записей в книгах, но пока отвечающий на вопрос не начал читать эти книги, в его наборе структур  $S$  отсутствуют необходимые компоненты  $M$  и  $R$ , при помощи которых описываются достаточно нетривиальные физические процессы.

Вышеописанные примеры описывают две типовых ситуации при постановке задач в виде вопроса, которые связаны с недостаточной компетентностью эксперта, однако имеющих существенную разницу в возможности своего разрешения в обозримый период времени. Другим хорошим примером некорректных вопросов первого рода является возможность создания рабочего прототипа коммерческого термоядерного реактора, на ответ на который понадобилось более пятидесяти лет и колоссальные инвестиции в проект ИТЕР. В данном случае, как и с вопросом о трансмутации, отсутствие необходимых открытий на момент постановки этой задачи является одной из причин её некорректности.

Примечательно, что причина некорректности столь нетривиальных задач в виде вопроса и вопросов типа «Существуют ли чётные трансфинитные числа?» одинакова и заключается в принципиальной невозможности ответа на эти вопросы в условиях мира  $W$ .

Формализованный язык<sup>1</sup>  $Я_{и1}$ , так же как и язык  $Я_{и}$ , представленный в работе Е.К. Войшвилло и Ю.А. Петрова, построен таким образом, что тривиально некорректные вопросы в нём невозможны, поскольку построение вопросов точно определено сформулированными правилами и все пропозициональные формулы, на основе которых создаются вопросы, а значит и сами вопросы, всегда осмысленны.

В то же время при постановке задачи в виде прогноза с использованием  $Я_{и}$  остаётся возможность постановки некорректных задач, так как они будут сформированы в виде некорректных вопросов второго рода.

Предположим, что в мире  $W$  имеется ЛПР, представления которого о мире  $W$  представляют собой фантомный мир  $W^{\setminus}_л$ . ЛПР знает о существовании некоего множества индивидов  $M_л$  и некоего множества отношений  $R_л$ , определённых на  $m$ -арном произведении множеств индивидов  $M_л$ , при этом  $W^{\setminus}_л \subseteq W$  и  $M_л$  и  $R_л \subseteq W^{\setminus}_л$ .

Также в мире  $W$  имеется Эксперт, призванный скорректировать постановку задачи, сформулированную ЛПР в виде вопроса. Представления Эксперта о мире  $W$  представляют собой фантомный мир  $W^{\setminus}_э$ . Эксперт знает о существовании некоего множества индивидов  $M_э$  и некоего множества отношений  $R_э$ ,  $z$ -арном произведении множеств индивидов  $M_э$ , при этом  $W^{\setminus}_э \subseteq W$  и  $M_э$  и  $R_э \subseteq W^{\setminus}_э$ . В данном случае множества  $M_л$  и  $M_э$ ,  $R_л$  и  $R_э$ , по сути, являются множествами представлений соответственно ЛПР и Эксперта о множествах  $M$  и  $R$ . Полагаем, что  $W^{\setminus}_э$  может пересекаться, а может и не пересекаться с миром  $W^{\setminus}_л$ .

Сформулируем эти особенности представлений ЛПР и Эксперта в виде правил интерпретации:

И1'. ЛПР каждой (интерпретируемой) предметной постоянной сопоставляется в качестве ее значения предмет из  $M_л$ .

---

<sup>1</sup> Язык  $Я_{и1}$  вполне можно считать формализованным языком, согласно определению: «формализованный язык в широком смысле – любая совокупность некоторым образом специализированных языковых средств с (более или менее) точно фиксированными правилами образования «выражений» (синтаксис Ф. Я.) и приписывания этим выражениям определённого смысла (семантика).» [8]. Соответственно, в  $Я_{и1}$  синтаксис определяется ПЯ<sub>и1</sub>–ПЯ<sub>и6</sub>, а смысл приписывается посредством И1–И4.

И1''). Эксперт каждой (интерпретируемой) предметной постоянной сопоставляется в качестве ее значения предмет из  $M_{\mathcal{E}}$ .

И2'. ЛПР каждой  $n$ -местной предикатной постоянной сопоставляется  $n$ -местное отношение из  $R_{\mathcal{L}}$ .

И2''). Эксперт каждой  $n$ -местной предикатной постоянной сопоставляется  $n$ -местное отношение из  $R_{\mathcal{E}}$ .

И3'. Для ЛПР областью значений предметных переменных является множество  $M_{\mathcal{L}}$ .

И3'') Для Эксперта областью значений предметных переменных является множество  $M_{\mathcal{E}}$ .

И4'. Для ЛПР предикатные переменные принимают значения из  $R_{\mathcal{L}}$  (причем значениями  $n$ -местной переменной могут быть лишь  $n$ -местные отношения).

И4''). Для Эксперта предикатные переменные принимают значения из  $R_{\mathcal{E}}$  (причем значениями  $n$ -местной переменной могут быть лишь  $n$ -местные отношения).

Также отметим, что поскольку в настоящей работе мы рассматриваем корректировку постановки некой управленческой задачи<sup>1</sup>, поставленной в виде вопроса, то вполне естественно ввести и использовать понятие множества целей ЛПР –  $P$ . Под множеством целей мы, опираясь на определение стратегической задачи из начала статьи, понимаем множество высказываний ЛПР о желаемом им состоянии фантомного мира  $W_{\mathcal{L}}^2$  являющегося совокупностью представлений ЛПР о реальном мире  $W$  в будущем, выраженных на языке  $Я_{i1}$ .

Теперь, закончив описание свойств формально-содержательного интеррогативного языка  $Я_{i1}$ , необходимых для доказательства утверждения о необходимых условиях для корректной постановки задачи в виде вопроса, сформулируем его и перейдем непосредственно к доказательству.

Вначале поясним понятие корректной формулировкой постановки управленческой задачи. Под корректной формулировкой постановки управленческой задачи при прогнозировании и планировании развития социально-экономических систем мы

---

<sup>1</sup> Которая подразумевает наличие у субъекта управления неких целей.

<sup>2</sup> Под состоянием мира здесь и далее понимается наличие в мире  $W_{\mathcal{L}}$  неких элементов множеств  $M_{\mathcal{L}}$  и  $R_{\mathcal{L}}$ , определённых ЛПР

будем понимать такую формулировку постановки задачи, которая удовлетворяет следующим двум требованиям:

II. Содержит в себе цели, которые ставит ЛПР при решении этой задачи

III. Такая формулировка должна быть понятна Эксперту, который будет решать эту задачу.

Поясним эти требования. В данном случае первое требование подразумевает, что наличие или отсутствие информации о целях ЛПР может вполне повлиять на выбор экспертом инструментария для её решения. А второе – что между представлениями Эксперта и ЛПР о предметной области, к которой относится данная задача, нет существенных расхождений.

Теперь сформулируем и докажем обоснованность этих требований в формальном виде.

Рассмотрим необходимые условия к корректной формулировке постановки задачи в виде вопроса, сформулированной на языке  $Я_{i1}$  в мире  $W$ , в виде следующего утверждения.

**Теорема 1. Необходимые условия корректной постановки задач в виде вопроса.**

**Вводная часть.** Пусть:

I.  $\{H\}$  – множество корректных постановок задач в виде вопроса, сформулированных на языке  $Я_{i1}$  в мире  $W$ , такое что  $\{H\} \subset W_{\text{Э}}$  и  $W_{\text{Л}}$ , где  $W_{\text{Э}}$  – фантомный мир эксперта, а  $W_{\text{Л}}$  – фантомный мир ЛПР.

II.  $\{Q_1\} \not\subset \{H\}$ , где  $\{Q_1\}$  – множество некорректных вопросов первого рода, сформулированных на языке  $Я_{i1}$  в мире  $W$ . При этом  $\{Q_1\} \subset \{Q\}$  и  $\{O_1\} \not\subset W$ , где  $\{Q\}$  – множество всех возможных вопросов, которые можно задать на языке  $Я_{i1}$  в мире  $W$ , а  $\{O_1\}$  – множество всех возможных ответов на вопросы из  $\{Q_1\}$ .

III. Пусть  $\{P\}$  – множество целей, а в данном контексте – высказываний ЛПР о желаемом им состоянии фантомного мира  $W_{\text{Л}}$ , являющегося совокупностью представлений ЛПР о реальном мире  $W$  в будущем, выраженных на языке  $Я_{i1}$ , такое, что  $\{H\} \cap \{P\} \subset W_{\text{Л}}$  и  $\{P\} \subset W_{\text{Л}}$ . Примем также еще одно допущение, что  $\{P\}$  не пусто. То есть множество корректных формулировок постановки задачи  $\{H\}$  включает в себя часть целей ЛПР

из множества  $\{P\}$ , которых он хочет достичь, решая данную задачу.

**Формулировка теоремы.** Докажем, что если Эксперту известны все цели ЛППР, которые контекстно связаны с рассматриваемой задачей, поставленной им же, и Эксперт знает обо всех объектах и отношениях между ними в мире  $W^1$ , о которых знает ЛППР, также контекстно связанных с задачей, то **любая** формулировка постановки задачи в виде вопроса, выраженная на языке  $Яи_1$ , корректна. В формальном виде последнее утверждение будет выглядеть следующим образом:

Если

IV.  $M_{л} \subseteq M_{э}$  и  $R_{л} \subseteq R_{э}$ ;

V.  $\{P\} \subseteq W^{\wedge}_{э}$ ;

то  $?A \subset \{H\}$ ,

где  $M_{л}$  – множество индивидов, о которых знает ЛППР;  $M_{э}$  – множество индивидов, о которых знает Эксперт;  $R_{л}$  – множество отношений, о которых знает ЛППР;  $R_{э}$  – множество отношений, о которых знает Эксперт;  $\{P\}$  – множество высказываний ЛППР о желаемом им состоянии фантомного мира  $W^{\wedge}_{л}$ , являющегося совокупностью представлений ЛППР о реальном мире  $W$  в будущем, выраженных на языке  $Яи_1$ ;  $W^{\wedge}_{э}$  – фантомный мир Эксперта;  $?$  – вопросный оператор;  $A$  – пропозициональная формула, содержащая более одной переменной; подразумевается, что  $A$  является формулировкой постановки задачи в виде вопроса.

**Доказательство.** В начале доказательства покажем целесообразность формулировки II, затем докажем, что  $I \& II \& III \rightarrow IV \& V$ , затем докажем необходимость IV, после этого покажем целесообразность формулировки III и докажем необходимость V.

Рассмотрим случай, указанный в II.

---

<sup>1</sup> Напомним, что в рамках данного исследования под миром  $W$  мы понимаем некий возможный мир, который является моделью производственной деятельности предприятия и его окружающей среды, в котором существуют объекты и отношения между ними, при этом одним объектом является ЛППР и хотя бы одним – Эксперт.

Предположим, что есть такая формулировка постановки задачи  $?A'' \in W_{\text{Э}}$  и  $W_{\text{Л}}$ , что ответ на неё  $O_{A''} \notin W$ . Тогда, исходя из правил И1' и И2'' языка Яи1, Эксперт будет составлять свой ответ из элементов множеств  $M_{\text{Э}}$  и  $R_{\text{Э}}$ , которые принадлежат  $W_{\text{Э}}$ . Поскольку  $W_{\text{Э}} \subseteq W$ , а  $O_{A''} \notin W$ , то Эксперт не сможет составить ответ на  $?A''$  при помощи языка Яи1. Тогда, поскольку  $W_{\text{Э}} \subseteq W$ , Эксперт не сможет сформулировать ответ на вопрос  $?A''$ .

Теперь покажем, что  $I \ \& \ II \ \& \ III \rightarrow IV \ \& \ V$ .

Рассмотрим каждое из условий левой части логического выражения по отдельности и покажем их необходимость.

Вначале рассмотрим случаи, обратные приведённому в I. Если  $\{H\} \not\subset W_{\text{Э}}$  и  $W_{\text{Л}}$ , тогда ни ЛПР, ни Эксперт не смогут сформулировать ни одну формулировку из  $\{H\}$ , поскольку Эксперт будет составлять свою интерпретацию формулировок из  $\{H\}$  из элементов множеств  $M_{\text{Э}}$  и  $R_{\text{Э}}$ , которые принадлежат  $W_{\text{Э}}$ , а ЛПР, соответственно, из элементов множеств  $M_{\text{Л}}$  и  $R_{\text{Л}}$ , которые принадлежат  $W_{\text{Л}}$ , тогда невозможно сформулировать никакую формулировку задачи из  $\{H\}$ .

В случаях когда  $\{H\} \not\subset W_{\text{Э}}$ , но  $\{H\} \subset W_{\text{Л}}$ , и  $\{H\} \subset W_{\text{Э}}$ , но  $\{H\} \not\subset W_{\text{Л}}$  соответственно, то Эксперт и ЛПР также не смогут сформулировать на языке Яи1 ни одну из формулировок постановки задач, входящих в  $\{H\}$ .

Таким образом,  $\{H\} \subset W_{\text{Л}} \cap W_{\text{Э}}$ .

Предположим, что  $W_{\text{Л}} \cap W_{\text{Э}}$  – это пересечение вида  $M_{\text{Л}} \subseteq M_{\text{Э}}$  и  $R_{\text{Л}} \subseteq R_{\text{Э}}$ , которое сформулировано в IV.

Теперь покажем необходимость условия II. Для задачи  $?A''$ , сформулированной в виде вопроса из  $\{Q_1\}$ , на который Эксперт не может дать ответ,<sup>1</sup> так как  $W_{\text{Э}} \subseteq W$ , а  $O_{A''} \notin W$ , к условию I необходимо добавить условие II. В противном случае может найтись формулировка постановки задачи  $Q^{**} \subset \{Q_1\}$ , а  $\{Q_1\} \cap \{P\}$  и  $Q^{**} \subset \{P\}$ , тогда, поскольку по условию III  $\{P\} \subset W_{\text{Л}}$ , а  $\{H\} \subset W_{\text{Л}} \cap W_{\text{Э}}$ , то  $Q^{**} \notin W_{\text{Л}} \cap W_{\text{Э}}$ , так как  $\{Q_1\} \not\subset \{H\}$ . То есть можно предположить, что если  $?A'' \in \{H\}$ , то  $?A'' \in W_{\text{Э}}$  и  $W_{\text{Л}}$ , и, следовательно,  $\{P\} \subseteq W_{\text{Э}}$ .

<sup>1</sup> В силу того, что этот вопрос выходит за рамки его знаний.

Поэтому можно предположить, что условие V следует из I & II & III.

Докажем необходимость условия IV. Для доказательства условия IV предположим обратное: пусть  $M_{\text{л}} \not\subseteq M_{\text{э}}$  и  $R_{\text{л}} \not\subseteq R_{\text{э}}$ , но  $?A \subset \{H\}$ . Тогда, исходя из правил интерпретации И1', И1'', И2' и И2'' любая постановка ЛПР задачи в виде вопроса ?A, сформулированная на языке Я<sub>и1</sub> и интерпретируемая ЛПР при помощи элементов множеств  $M_{\text{л}}$  и  $R_{\text{л}}$ , будет интерпретирована Экспертом при помощи элементов множеств  $M_{\text{э}}$  и  $R_{\text{э}}$ .

Тогда всегда, когда  $M_{\text{л}} \neq \emptyset$ , из-за того что в описании языка Я<sub>и1</sub>, указанного выше, фантомный мир  $W_{\text{э}}$  может пересекаться, а может и не пересекаться с миром  $W_{\text{л}}$ , найдётся хотя бы одна формулировка ?A', интерпретация которой будет принадлежать  $W_{\text{э}}$ , но не будет принадлежать  $W_{\text{л}}$ , а поскольку  $M_{\text{л}} \subseteq W_{\text{л}}$ , то такая интерпретация ?A не принадлежит H, следовательно формулировка постановки задачи ?A  $\not\subseteq H$  и потому не может быть приведена к корректному виду.

Для случаев  $M_{\text{л}} \not\subseteq M_{\text{э}}$  и  $R_{\text{л}} \subseteq R_{\text{э}}$  доказательство аналогично.

Рассмотрим случай, когда  $M_{\text{л}} \subseteq M_{\text{э}}$  и  $R_{\text{л}} \not\subseteq R_{\text{э}}$ . Тогда, исходя из правил интерпретации И1 и И2 и описанию поведения Эксперта, любая постановка ЛПР задачи в виде вопроса ?A, сформулированная на языке Я<sub>и1</sub> и интерпретируемая ЛПР при помощи элементов множеств  $M_{\text{л}}$  и  $R_{\text{л}}$ , будет интерпретирована Экспертом при помощи элементов множеств  $M_{\text{э}}$  и  $R_{\text{э}}$ . Тогда всегда, когда  $R_{\text{л}} \neq \emptyset$ , найдётся хотя бы одна корректированная формулировка ?A, интерпретация которой будет принадлежать  $W_{\text{э}}$ , но не будет принадлежать  $W_{\text{л}}$ , а поскольку  $R_{\text{л}} \subseteq W_{\text{л}}$ , то такая интерпретация ?A не принадлежит {H}, следовательно формулировка постановки задачи ?A  $\notin H$  и потому не может быть приведена к корректному виду.

Рассмотрим случай, указанный в III.

Как было показано в доказательстве условия IV,  $W_{\text{л}} \cap W_{\text{э}}$  в {H}, а именно,  $M_{\text{л}} \subseteq M_{\text{э}}$  и  $R_{\text{л}} \subseteq R_{\text{э}}$ . Тогда поскольку {P} есть множество высказываний ЛПР, которые можно выразить при помощи  $M_{\text{л}}$  и  $R_{\text{л}}$ , то можно предположить, что  $\{P\} \subseteq W_{\text{э}}$ .

Докажем необходимость условия V.

Предположим обратное указанному в III, что  $\{P\} \notin W_{\text{Э}}$ . Исходя из правил интерпретации  $I1'$ ,  $I1''$ ,  $I2'$  и  $I2''$ , любая постановка ЛПР задачи в виде вопроса ?А, сформулированная на языке  $Я_{i1}$  и интерпретируемая ЛПР при помощи элементов множеств  $M_{\text{Л}}$  и  $R_{\text{Л}}$ , будет интерпретирована Экспертом при помощи элементов множеств  $M_{\text{Э}}$  и  $R_{\text{Э}}$ .

Тогда при корректировке постановки задачи, сформулированной ЛПР в виде вопроса ?С, всегда найдётся хотя бы одна скорректированная формулировка постановки задачи ?С, такая что ни одно высказывание ? $p \notin ?С$ . Поскольку при интерпретации Экспертом постановки задачи ?С содержащиеся в ней ? $p$  не могут быть интерпретированы при помощи множеств  $M_{\text{Э}}$ ,  $R_{\text{Э}}$ , так как  $\{P\} \notin W_{\text{Э}}$ , то при интерпретации и корректировке полученной формулировки постановки задачи в виде вопроса ?С всегда может быть получена такая ?С, что ? $p \notin ?С$  не может принадлежать  $H$ . Следовательно, такая корректировка постановки задачи ?С не может быть приведена к корректному виду.

Таким образом, при невыполнении условий IV и V формулировка постановки задачи в виде вопроса не может быть приведена к корректному виду. Утверждение доказано.

Формализованный язык  $Я_{i1}$ , представленный в данной работе, построен таким образом, что нетривиально некорректные вопросы в нём невозможны, поскольку построение вопросов точно определено сформулированными правилами и все пропозициональные формулы, на основе которых создаются вопросы, а значит и сами вопросы, всегда осмысленны.

Таким образом, язык  $Я_{i1}$  расширяет возможности языка  $Я_{i0}$  по корректировке постановки задач в виде вопроса. Основными нововведениями, отличающими язык  $Я_{i1}$  от языка  $Я_{i0}$ , являются:

1. Рассмотрение возможности различной интерпретации высказываний в зависимости от типа агента (ЛПР или Эксперт).
2. Наличие формализованных определений некорректных вопросов.
3. Требования к ответам, снижающие количество ответов близких по смыслу, но различающихся по форме.



Рассматриваемые нововведения позволяют сформулировать формальное определение корректной постановки задачи и необходимые условия для приведения некорректной формулировки постановки задачи к корректной.

Используя язык Я<sub>1</sub>, базирующийся на языке Войшвилло – Петрова, для формализации формулировок постановок задачи в виде вопроса средствами естественного языка, на выходе мы получим некоторое количество корректных вопросов, на которые в принципе возможен ответ, при этом будет отсеяно большое количество некорректных вопросов и таким образом будет сэкономлено значительное количество времени, которое в противном случае было бы потрачено на некоторые уточняющие вопросы. Используя язык Я<sub>1</sub> можно отсеять как откровенно ошибочные вопросы, вроде «Как повысить ставку разфондирования?», так и вопросы, содержащие неочевидную некорректность например: «Будет ли полезно снизить ставки по ипотеке?». И если эффект от отбраковки откровенно ошибочных вопросов относительно невелик, то в случае неочевидной некорректности вопроса наличие формального критерия некорректности задачи позволяет экономить время и ресурсы, которые иначе были бы потрачены впустую на решение некорректной задачи.

В случае с вопросом о снижении ипотечных ставок некорректность вопроса заключается в отсутствии в формулировке задачи упоминания целей ЛПР. И в данном случае ответ эксперта будет зависеть от того, для какого участника рынка эксперт будет оценивать полезность. Если речь идёт о застройщике и банке, то для них это будет чрезвычайно полезно, поскольку цены на недвижимость вырастут, как и их выручка. Покупателя ипотеки снижение стоимости займа скорее поставит в проигрышное положение из-за роста суммарных затрат на покупку недвижимости из-за роста его цены. Государство в таком случае также в выгоде, поскольку рост цен на недвижимость ведёт к росту выручки компаний строительного сектора и части сырьевых компаний, которые обеспечивают сохранение рабочих мест.

Оценка полезности снижения ставок для всех участников рынка требует значительного времени и ресурсов, в то время как

для ЛПР может быть важна полезность только для одного из участников рынка. Экономия времени и ресурсов также имеет свои плюсы для крупных компаний, где повторное согласование задач после выявления некорректности первоначальных прекращается в довольно продолжительную бюрократическую процедуру. По сути, все ранее подготовленные документы проходят согласование повторно.

Таким образом, за счёт формализации вопроса с учётом имеющейся и очевидно известной спрашивающему посылки, можно дать относительно понятный отвечающему на вопрос критерий для отбора ответа из множества имеющихся альтернатив.

## **5. Выводы**

В данной работе рассматривалась проблема уменьшения влияния негативных факторов на корректность постановки задачи. Поскольку эти факторы обладают сложной, комплексной природой, в работе был представлен подход, основанный на использовании интеррогативной логики, который, по мнению автора, способен уменьшить степень влияния этих факторов на эффективность принятия управленческих решений.

В рамках создания данного подхода был решён ряд задач. Для решения проблем, связанных с корректностью постановки управленческих задач при активном прогнозировании, был разработан механизм корректировки постановки управленческих задач для информационного управления социально-экономическими системами, основанный на аппарате интеррогативной логики.

Был сформулирован и обоснован список необходимых условий для корректной постановки управленческой задачи при информационном управлении социально-экономическими системами при помощи активных прогнозов.

Также был сформулирован специальный формально-содержательный интеррогативный язык, использующийся для уменьшения потерь информации при передаче задания исполнителю.

## Литература

1. АЛЕКСАНДРОВ А.А. *Интегративная психотерапия*. — СПб.: Питер, 2009. — 352 с.
2. БЕЛНАП Н., СТИЛ Т. *Логика вопросов и ответов*. — М.: Прогресс, 1981. — 290 с.
3. БЕРКОВ В.Ф. *Вопрос как форма мысли*. — Мн., 1972.
4. БОЧАРОВ В.А., МАРКИН В.И. *Введение в логику: учебник*. — М.: ИД «ФОРУМ»: ИНФРА-М, 2008. — 560 с.
5. ВОЙШВИЛЛО Е.К., ПЕТРОВ Ю.А. *«Язык и логика вопросов»* // В сб.: «Логика и методология научного познания» / Под ред А.А. Старченко. — М.: Изд. московского университета, 1974. — С. 147–158.
6. СИДЕЛЬНИКОВ Ю.В., САЛТЫКОВ С.А. *Процедура установления соответствия между задачей и методом* // Экономические стратегии. — 2008. — №7(65). — С. 102–109.
7. ФЕДОРОВ Б.И. *«О классификации научных вопросов»*, [Электронный ресурс] [http://ecsocman.hse.ru/data/2010/11/02/1214794744/%D0%A4%D0%B5%D0%B4%D0%BE%D1%80%D0%BE%D0%B2\\_%D0%9F%D1%80%D0%BE%D0%B1%D0%BB%D0%B5%D0%BC%D0%B0\\_%D0%BA%D0%BB%D0%B0%D1%81%D1%81%D0%B8%D1%84%D0%B8%D0%BA%D0%B0%D1%86%D0%B8%D0%B8\\_%D0%B2%D0%BE%D0%BF%D1%80%D0%BE%D1%81%D0%BE%D0%B2.pdf](http://ecsocman.hse.ru/data/2010/11/02/1214794744/%D0%A4%D0%B5%D0%B4%D0%BE%D1%80%D0%BE%D0%B2_%D0%9F%D1%80%D0%BE%D0%B1%D0%BB%D0%B5%D0%BC%D0%B0_%D0%BA%D0%BB%D0%B0%D1%81%D1%81%D0%B8%D1%84%D0%B8%D0%BA%D0%B0%D1%86%D0%B8%D0%B8_%D0%B2%D0%BE%D0%BF%D1%80%D0%BE%D1%81%D0%BE%D0%B2.pdf) [дата обращения: 01.04.2021].
8. ЧЁРЧ А. *Введение в математическую логику*. Том 1. — М.: Изд-во иностранной литературы, 1960. — 484 с.
9. AHLEMANN F., STETTINER E., MESSERSCHMIDT M., LEGNER C. *Strategic Enterprise Architecture Management: Challenges, Best Practices, and Future Developments*. — Springer-Verlag Berlin Heidelberg, 2012. — 300 p.
10. BAER M., DIRKS K.T., NICKERSON J.A. *Microfoundations of strategic problem formulation* // Strategic Management Journal. — 2013. — Iss. 34. — P. 197–214.
11. EISENHARDT K.M., MARTIN J.A. *Dynamic capabilities: what are they?* // Strategic Management Journal. — 2000, October–November. — Special Issue 21. — P. 1105–1121.

12. HAMAMI Y. *The interrogative model of inquiry meets dynamic epistemic logics* // *Synthese*. – 2015. – Vol. 192(6). – P. 1609–1642.
13. HAMBLIN C.L. *Questions* // *Australasian Journal of Philosophy*. 1958. – Vol. 36. – P. 158–159.
14. HARRAH D. *A logic of questions and answers* // *Philosophy of Science*. – 1961. – Vol. 28. – P. 40–46.
15. HINTIKKA J. *Inquiry as inquiry: A logic of scientific discovery*. – Kluwer Academic Publishers, 1999. – 304 p.
16. HINTIKKA J. *Knowledge and belief. An introduction to the logic of the two notions*. – Cornell University Press, Ithaca, N.Y., 1962. – 179 p.
17. HINTIKKA J. *Socratic epistemology: explorations of knowledge-seeking by questioning*. – Cambridge University Press, 2007. – 248 p.
18. HOFSTEDE G., MINKOV M., *Organizations: Software of the Mind*. – New York: McGraw-Hill USA, 2010. – 550 p.
19. KUBINSKI T. *An Essay in Logic of Questions* // In: «Atti del XII Congresso Inter, do Filosofia», Venezia, 1958. – Bd. 5.
20. MARITAN C.A. *Dynamic capabilities and organizational processes* // In: «Dynamic Capabilities: Understanding Strategic Change in Organizations» /Eds.: Helfat C.E., Finkelstein S, Mitchell W, Peteraf M.A., Singh H., Teece D.J., Winter S.G. – Blackwell Publishing: Malden, MA, 2007. – P. 30–45.
21. PRIOR A., PRIOR M. *Erotetic logic* // *Philosophical Review*. – 1955. – Vol. 64. – P. 43–59.
22. SHRIVASTAVA P, GRANT J.H. *Empirically derived models of strategic decision-making processes* // *Strategic Management Journal*. – 1985. – Vol. 6(2). – P. 97–113.
23. SPERANTIA E. *Remarques sur les propositions interrogatives. Projet d'une "logique du problème"* // In: «Actes du Congrès International de Philosophie Scientifique», VII Logique, Paris 1936. – P. 18–28.
24. WHATELY R. *Elements of Logic*. – Longman, Greens and Co., 1875.
25. WISNIEWSKI A. *Erotetic Search Scenarios, Problem-Solving, and Deduction* // *Logique & Analyse*. – 2004. – No. 185–188. – P. 139–166.

26. *A cost-benefit analysis of the Concorde Project.* – [Электронный ресурс] <https://www.jstor.org/stable/20052287?seq=1> [дата обращения: 08.02.2021].
27. *A39 Tortoise Heavy Assault Tank.* – [Электронный ресурс] <http://www.armourinfocus.co.uk/a39/a39txt.htm> [дата обращения: 08.02.2021]
28. *How to set strategic planning goals.* – [Электронный ресурс] <https://online.hbs.edu/blog/post/strategic-planning-goals> [дата обращения: 08.02.2021].

## CORRECTNESS OF CONTROL TASK SETTING

**Maxim Shevyrenkov**, postgraduate student, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Russian Federation (sheverenkov@mail.ru).

*In this paper, the problem of reducing the influence of negative factors on the correctness of the problem statement was considered. Since these factors are of a complex, complex nature, this paper presented an approach based on the use of interrogative logic, which, according to the author, is able to reduce the degree of influence of these factors on the effectiveness of managerial decision-making. As part of the creation of this approach, a number of problems were solved. To solve the problems associated with the correctness of the formulation of managerial tasks with active forecasting, a mechanism for adjusting the formulation of managerial tasks for information management of socio-economic systems was developed based on the apparatus of interrogative logic. A list of necessary conditions for the correct formulation of a management problem in the information management of socio-economic systems using active forecasts was formulated and substantiated. Also, a special formal-meaningful interrogative language was formulated, which is used to reduce the loss of information when transferring the task to the performer.*

Keywords: interrogative logic, correctness of task setting, control tasks, strategic management, strategic planning, decision making, management task.

УДК 007

ББК 65.05

DOI: 10.25728/ubs.2021.90.4

*Статья представлена к публикации  
членом редакционной коллегии Ю.В. Сидельниковым.*

*Поступила в редакцию 30.11.2020.*

*Опубликована 31.03.2021.*

## ИНТЕГРИРОВАННЫЕ АВИАЦИОННЫЕ СИСТЕМЫ

Клочков В. В.<sup>1а,б</sup>, Топоров Н. Б.<sup>2б,в</sup>, Егوشин С. Ф.<sup>3б</sup>

<sup>а</sup> ФГБУН Институт проблем управления

им. В.А. Трапезникова РАН, Москва,

<sup>б</sup> ФГБУ НИЦ «Институт им. Н.Е. Жуковского»,  
Жуковский,

<sup>в</sup> ФГБОУ ВО «Московский авиационный институт (национальный исследовательский университет)», Москва)

*Повышение эффективности управления научно-технологическим развитием авиастроения на данном этапе возможно за счет перехода к рассмотрению крупномасштабных систем с большим количеством «степеней свободы». Впервые предложен и развит методологический подход к прогнозированию и стратегическому планированию научно-технологического развития авиастроения, основанный на анализе и синтезе не отдельных образцов авиационной техники или авиационных комплексов, а интегрированных систем. Эти системы направлены на решение целевых задач высокого уровня и включают в себя парки различных воздушных судов, а также инфраструктурные и управляющие подсистемы, включая неавиационные подсистемы целевого заказчика. Методология исследования базируется на теории систем, методах системного анализа и синтеза, математическом моделировании. Предложенный подход позволяет выявлять потенциальные области применения перспективной авиационной техники в сфере транспорта и различных отраслях национальной экономики, а также оценивать потенциальные масштабы применения авиации. Результаты моделирования и оптимизации интегрированных авиационных систем должны стать основой планирования приоритетных научных исследований и разработок в авиастроении, в том числе при формировании государственных программ Российской Федерации, стратегий развития научных и промышленных организаций авиастроения.*

**Ключевые слова:** интегрированная авиационная система, моделирование, оптимизация, проектирование летательных аппаратов, системный анализ, научно-технологическое развитие авиастроения.

---

<sup>1</sup> Владислав Валерьевич Клочков, д.э.н. (klochkovvv@nrczh.ru).

<sup>2</sup> Николай Борисович Топоров, д.т.н., профессор (toporov@nrczh.ru).

<sup>3</sup> Сергей Федорович Егوشин, специалист (sergey4791@yandex.ru).

## **1. Введение**

Развитие математических методов и их практическая реализация с применением компьютерной техники с учетом ее совершенствования позволили разработчикам авиационной техники начиная со второй половины XX века постепенно повышать уровень сложности моделирования [17, 18] и оптимизации как летательных аппаратов (ЛА) в отдельности, так и среды их применения. Системные исследования стали значимой областью прикладной авиационной науки [3, 4, 10, 14]. Не очерчивая строгие временные рамки, можно условно выделить следующие этапы развития этой области.

1. Оптимизация ЛА (здесь и далее – одно- или многокритериальная).

Математическая постановка задачи сводится к отысканию наилучшего сочетания проектных параметров, при которых реализуется оптимум целевой функции (если критерий один) или множество эффективных вариантов (множество Парето для набора частных критериев эффективности, если таковых критериев более одного).

2. Оптимизация обликов авиационных комплексов (АК).

Так как АК – функционально связанная система, включающая летательный аппарат, бортовое оборудование и целевое оснащение, экипаж, средства обеспечения и обслуживания специального назначения и предназначенная для выполнения некоторого множества задач [6, 12], то процесс оптимизации параметров облика АК от предыдущего случая отличается тем, что рассматриваемый объект моделирования – это не обособленный ЛА, а ЛА в сочетании с оборудованием/снаряжением, которое может создаваться независимо от ЛА. Так, например, типичным образцом подобного АК может служить комплекс ПВО, состоящий из истребителя-перехватчика как носителя ракет класса «воздух – воздух», собственно указанных ракет, комплекса радиосвязного оборудования для целеуказания с земли и единой системы управления и наведения. Оптимизация облика АК (т.е. решение задачи синтеза оптимального облика АК) позволяет

повысить ожидаемую эффективность применения разработанного АК.

### **3. Оптимизация обликов авиационных систем (АС).**

В данном случае под АС следует понимать совокупность АК (возможно, разнотипных по назначению) и обеспечивающей их жизнедеятельность авиационной инфраструктуры. Совместное использование АК и инфраструктуры позволяет решить определенный набор задач [11, 16]. Таким образом, в АС, помимо собственно АК, также входят аэропорты и посадочные площадки, топливозаправочные комплексы и склады ГСМ, объекты системы организации воздушного движения (ОрВД), авиаремонтные заводы (АРЗ), склады запасных частей, сменных узлов и агрегатов, и прочие производственные, сервисно-логистические и управляющие подсистемы.

Однако в настоящее время, с одной стороны, назрела необходимость, а с другой стороны – появились возможности рассмотрения и оптимального проектирования систем еще более высокого уровня [13, 15, 19, 20], включающих в себя как собственно АС, так и «ответную часть», объекты целевой области применения авиации. В данной работе обоснована актуальность постановки и решения таких задач, предложена формализация их постановки.

## **2. Проблемы моделирования и оптимизации авиационных систем различного уровня**

Исследование АС является комплексным подходом, позволяющим найти более предпочтительные управляющие и проектные решения, в отличие от оптимизации АК и, тем более, ЛА. «Платой» за возможность получения этого решения становятся: проблема рационального выбора критериальной базы решения задачи, значительное усложнение самого процесса поиска решения задачи (вследствие увеличения ее размерности) и повышение рисков принятия некорректного решения.

Так, например, ошибки при решении задачи прогнозирования спроса на перевозки пассажиров в масштабах мировой авиатранспортной системы, АТС (которая также относится



## Управление техническими системами и технологическими процессами

к задаче оптимизации АС) привели компанию Airbus к решению создать самолет сверхбольшой пассажироместимости А380. Однако дальнейшее производство и эксплуатация А380 показали, что, несмотря на общий рост объемов авиационных пассажироперевозок, самолет пользуется ограниченным спросом и его производство – убыточно.

В графической форме иерархию описанных уровней моделирования и оптимизации систем в области авиации и авиастроения можно представить в следующем виде:

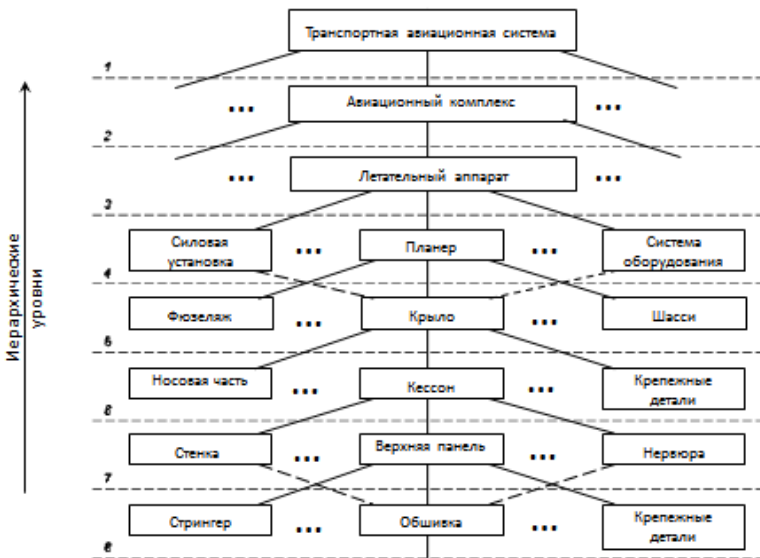


Рис. 1. Иерархия структуры летательного аппарата как части большой технической системы [5]

Уровни 1–3 можно отнести к этапу внешнего проектирования АК (ЛА), и каждому из этих уровней соответствует один из трех вышеуказанных типов оптимизации. При этом проектирование АС, соответствующее верхнему, первому уровню, позволяет не только сформировать облик ЛА, но и определить роль

и место АК (ЛА) в общем парке, потребное их количество, требования к аэродромам базирования, составу экипажей и наземного персонала и т.д.

Уровни же 4–8 соответствуют задачам, решаемым либо авиационной наукой в кооперации с разработчиками ЛА или АК, либо непосредственно в конструкторских бюро (формирование технического предложения, аванпроекта и эскизного проекта, технической документации).

Для уровней 1–3 математическую формулировку задачи оптимизации можно представить в следующем виде:

$$(1) \begin{cases} \text{extr } S(X), \\ F(X) \geq A, \\ G(X) \leq B; \end{cases}$$

где  $S$  – векторная целевая функция (критерий оптимальности);  $X = \{x_1, \dots, x_m\}$  – вектор проектных параметров оптимизируемой системы;  $F(X)$  – векторная функция требуемых уровней (объемов) выполнения  $P$  целевых задач  $A$  ( $A = \{a_1, \dots, a_n\}$ );  $G(X)$  – векторная функция  $Q$  ограничений  $B$  ( $B = \{b_1, \dots, b_q\}$ ).

Для каждой из технических систем, соответствующих уровням 1–3, входящие в постановку задачи (1) величины и функции будут различаться. При этом для системы более высокого уровня параметрами (управляющими переменными) могут являться как характеристики (свойства), так и параметры системы более низкого уровня.

В стремлении найти глобальный оптимум АТС целесообразна такая постановка задачи оптимизации, в которой вектором проектных параметров  $X$  этой системы верхнего уровня будут являться выходные параметры систем более низкого уровня.

Но тогда при каждом понижении уровня (т.е. детализации, разукрупнении) системы, параметры которой будут использоваться в качестве вектора  $X$  системы 1-го уровня, будет происходить значительное усложнение постановки задачи 1-го уровня (и соответственно усложняться отыскание решения).

То есть «сквозная» оптимизация при проектировании подобной многоуровневой системы «в один проход» как в направлении «снизу вверх», так и «сверху вниз» на практике обречена

## *Управление техническими системами и технологическими процессами*

на неудачу. Это вызвано не только катастрофическим ростом размерности задачи. При попытках ее декомпозиции по уровням иерархии, если в системе проектирование происходит по принципу «снизу вверх» (решение, по какому-либо критерию оптимальное для системы нижнего уровня, становится входными, проектными параметрами для системы вышерасположенного уровня), то найденный оптимум системы верхнего уровня далеко не всегда будет соответствовать глобальному оптимуму системы. Если же проектирование вести «сверху вниз», оптимальное решение задачи верхнего уровня может оказаться нереализуемым, недопустимым для систем нижнего уровня.

Приведенное противоречие можно проиллюстрировать примером задачи оптимизации АТС местных авиалиний в отдаленных, труднодоступных и малонаселенных районах (ОТДМР) Российской Федерации. Общее количество аэродромов и посадочных площадок, необходимых для обслуживания ОТДМР, составляет несколько сотен. В силу социально-экономических факторов местные авиаперевозки в России не являются экономически прибыльными, а потому требуют дотирования со стороны государства. Следовательно, критерием оптимальности  $S$  на верхнем уровне иерархии решения задачи целесообразно считать затраты государства на поддержание некоторого уровня транспортного обслуживания ОТДМР (например, поддержание заданного уровня авиационной мобильности населения этих регионов при заданной частоте рейсов или даже при заданном гарантированной времени в пути «от двери до двери»). Тогда если в качестве исходных данных принимаются характеристики существующих типов ВС «малой авиации», то задача (1), решаемая в постановке верхнего 1-го уровня, может быть сведена к задаче целочисленного программирования. В этой задаче вектор  $X$  будет состоять из управляющих булевых переменных, включающих или исключающих задействование каждого из аэродромов, а также целочисленных переменных, определяющих количество рейсов для каждой возможной авиалинии и состав потребного парка ВС.

Современные математические методы позволяют найти точное решение приведенной целочисленной задачи. Вместе

с тем полученное решение не обязательно будет находиться в окрестности глобального оптимума затрат системы, в первую очередь ввиду ограниченности альтернативных вариантов ВС (только существующие типы), среди которых может и не быть оптимальных. Более того, на практике полученное на основе существующего типажа ВС решение может оказаться недопустимым либо потому что такая АТС обеспечит недопустимо низкое качество авиатранспортного обслуживания, очень редкие рейсы, либо потому что потребные затраты государства на поддержание заданного транспортного стандарта выйдут за рамки реальных возможностей бюджета.

Чтобы найти более предпочтительное решение, в постановке задачи верхнего, 1-го уровня целесообразно заменить набор априорно заданных типов ВС на математическую модель произвольного самолета местных воздушных линий (так называемый «резинový самолет»), варьируемые параметры которой дополнительно войдут в вектор  $X$ . Данный шаг приведет к тому, что задача из целочисленной превратится в непрерывную, причем существенно нелинейную. Но поскольку возросшие возможности современной компьютерной техники позволяют найти приближенное решение для многих подобных задач применительно к большим техническим системам, это открывает возможности для дальнейшего усложнения исходной задачи 1-го уровня.

При этом рост сложности возможен не только за счет расположенных ниже уровней иерархии (т.е. с добавлением в расчетную модель переменных, служащих для описания моделей 4-го, 5-го и т.д. уровней), но и за счет расширения состава рассматриваемой надсистемы. Предпосылки и возможности для этого шага излагаются ниже.

В качестве основного подхода в этой ситуации может быть предложена циклическая двунаправленная процедура поиска рационального решения, объединяющая каждую пару уровней последовательно как при движении вверх, так и вниз. Причем начало такой последовательности опирается на наиболее высокий в иерархии уровень, который и определяет как критериальную базу с дальнейшей декомпозицией критериев для более

низких уровней решения задачи, так и последовательность ее решения [11].

### **3. Предпосылки перехода к рассмотрению интегрированных авиационных систем**

В XX веке авиация достаточно быстро сформировалась именно как сложная организационно-техническая система. Ее возникновение и развитие было обусловлено не только возможностями научно-технического прогресса, но и интересами и нуждами граждан, социальных групп, предприятий, государств. Авиация стала не самоцелью, как на этапе ее становления, а средством (причем, как правило, лишь одним из многих альтернативных) решения различных проблем общества и государства, вплоть до уровня национальной и мировой экономики [2].

Однако авиация зачастую не является единственным способом решения этих проблем, а чаще одним из возможных инструментов. И далеко не обязательно какие-то целевые задачи в народном хозяйстве будут решаться именно привычными авиационными средствами, и наоборот, авиация может найти новые области применения. Поэтому при решении задачи оптимального построения АС необходимо учитывать все компоненты, включая не только обеспечивающие функционирование авиации (авиационную инфраструктуру), но и решающие целевые задачи в единой надсистеме более высокого уровня (т.е. уже неавиационную, «целевую» инфраструктуру).

Таким образом, решающая роль в повышении эффективности авиационной деятельности переходит уже от оптимизации АС к оптимизации всей среды (сферы деятельности или отрасли экономики), в которой применяется авиация.

В качестве примера может быть приведена санитарная авиация как «длинная рука» системы здравоохранения [7]. Наибольший эффект от ее развития будет получен отнюдь не при локальной оптимизации парка ВС, аэродромов и посадочных площадок для фиксированной сети обслуживаемых медицинских учреждений. Он будет получен именно при глобальной оптимизации пространственной организации здравоохранения,

включая расположение, масштаб и специализацию медицинских учреждений с учетом возможностей, предоставляемых санитарной авиацией.

При этом и система медицинских учреждений, и санитарная авиация должны быть многоуровневыми, эшелонированными, в зависимости от сложности соответствующих видов медицинской помощи и интенсивности соответствующих «заявок». Так, первую медицинскую помощь целесообразно оказывать на базе фельдшерско-акушерских пунктов, и в несложных случаях, когда требуется консультация профильных специалистов, – прибегая к технологиям телемедицины. В более сложных случаях, когда необходимо личное посещение профильного специалиста или оказание высокотехнологичной медицинской помощи, доступ к соответствующим медицинским учреждениям возможен при организации санитарных перевозок (в том числе авиационных) между населёнными пунктами и медицинскими учреждениями. Для оказания неотложной медицинской помощи и эвакуации пострадавших в тяжёлом и критическом состоянии целесообразно использовать именно санитарную авиацию. ЛА санитарной авиации также могут быть разнотипными. Часть из них должна обладать возможностью использования неподготовленных взлётно-посадочных площадок, другие санитарные ЛА должны выполнять скоростные дальние перевозки с использованием аэродромов. Но важнее, что вся система в целом – «сеть медучреждений плюс санитарная авиация» – должна обеспечивать необходимую оперативность медицинской помощи.

Таким образом, из понимания роли авиации как части соответствующей отрасли народного хозяйства и следует необходимость разработки новых подходов к прогнозированию и планированию ее развития.

Наиболее системный подход, позволяющий получать глобально оптимальные решения, состоит в рассмотрении (анализе путем математического и компьютерного моделирования и последующем оптимальном синтезе) интегрированных авиационных систем (ИАС), включающих в себя:

- многотипный парк ВС различного назначения;

## Управление техническими системами и технологическими процессами

- объекты авиационной инфраструктуры;
- неавиационные объекты отраслей народного хозяйства (промышленные и сельскохозяйственные предприятия, медицинские учреждения, государственные территориальные службы и т.д.), которые и выполняют основную целевую функцию и для обеспечения деятельности которых и предназначена рассматриваемая специализированная ИАС.

Первые два компонента из трех, как было сказано выше, составляют АС, добавление третьего превращает ее в ИАС:

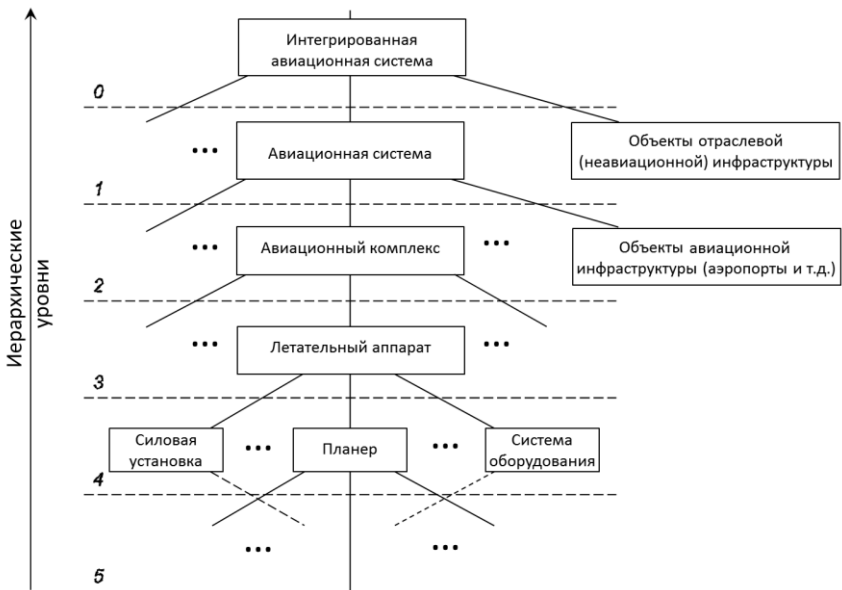


Рис. 2. Интегрированная авиационная система

Отсюда можно дать следующее определение: ИАС – это совокупность различных авиационных комплексов (воздушных судов), наземной и воздушной управляющих, обеспечивающих и целевых инфраструктур, совместно и координированно решающая определенный спектр целевых задач в условиях изменяющегося внешнего окружения и при соответствующих ограничениях.

В состав ИАС может входить несколько АС. Однако ключевым фактором, отличающим АС от ИАС, является включение в систему объектов инфраструктуры, непосредственно решающей целевые задачи. Условное обозначение уровня «0» на рис. 2 следует трактовать не только как продолжение нумерации рисунка 1 (во избежание путаницы). Тем самым мы также косвенно подчеркиваем тот факт, что еще более вышестоящий уровень анализа и оптимизации на данном этапе развития авиационной науки сложно сформулировать.

Различные ИАС могут развиваться почти изолированно друг от друга и рассматриваться в отдельности. В то же время различные ИАС могут и объединяться на уровне отдельных подсистем. Их общими структурными элементами могут быть, например, аэродромы, объекты ОрВД или предприятия авиационно-промышленного комплекса (его предприятия работают в интересах многих ИАС).

#### **4. Обоснование целесообразности рассмотрения интегрированных авиационных систем**

Плодотворность моделирования ИАС (и решения соответствующей оптимизационной задачи самого верхнего уровня) может быть проиллюстрирована следующим образом.

Пусть целевой критерий оптимизации  $C_{\Sigma}$  – это стоимость создания и применения ИАС в рассматриваемый период планирования. Суммарные затраты  $C_{\Sigma}$  складываются из затрат  $C_{\text{АВИА}}$  на АС, являющуюся компонентой системы, и непосредственно отраслевых затрат  $C_{\text{ОТР}}$ , приходящихся на остальные, неавиационные компоненты целевой области применения – отрасли, предприятия и т.п. В примере санитарной авиации это сеть учреждений. Если инфраструктура отрасли рассматривается как фиксированная (не подлежащая оптимизации), можно, разве что, найти оптимум затрат на АС (при заданном объеме авиационных работ и услуг, который обозначается величиной  $X$  и отображается по оси абсцисс), равный  $C_{\text{АВИА}}^0$ , и соответствующие ему отраслевые затраты  $C_{\text{ОТР}}^0$ , см. рис. 3.



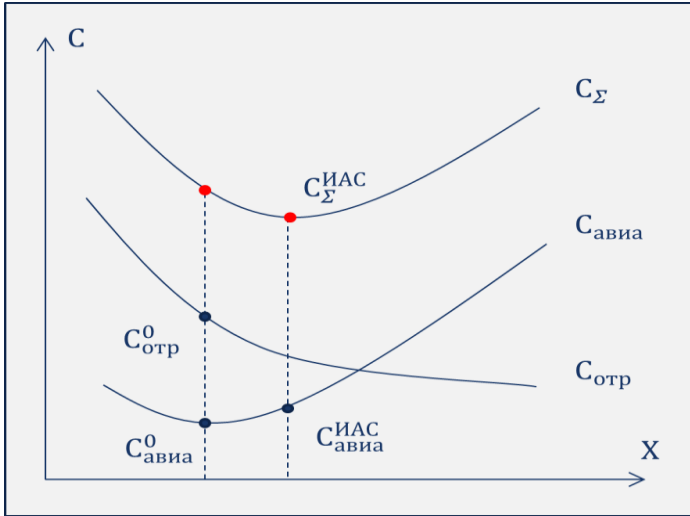


Рис. 3. Условие существования ИАС

Однако если оптимизация затрат производится в рамках единой постановки задачи оптимизации ИАС (т.е. в которой АС не рассматривается отдельно, а процесс оптимизации затрагивает все, в том числе неавиационные компоненты системы), то это, в зависимости от условий, может обеспечить более или менее значительное снижение суммарных затрат:

$$(2) \quad C_{\Sigma}^{\text{ИАС}} < C_{\text{АВИА}}^0 + C_{\text{ОТР}}^0.$$

Приведенный график показывает наглядно, что интересы заказчика не всегда сводятся лишь к тому, чтобы получить определенные авиационные работы или услуги с наименьшими затратами. Его целью является наиболее эффективное удовлетворение своих конечных потребностей, теми или иными средствами. Поэтому вполне возможно, что, например, при снижении цены и/или повышении качества авиационных работ и услуг оптимальный спрос на них возрастет настолько, что общие затраты заказчиков именно на авиационные работы и услуги увеличатся:

$$(3) \quad C_{\text{АВИА}}^{\text{ИАС}} > C_{\text{АВИА}}^0.$$

И, тем не менее, это будет эффективно для конечного заказчика, для области целевого применения авиации.

Таким образом, проблема поиска рационального объема применения авиации (и, как следствие, – оценка спроса на продукцию авиастроения) в первую очередь относится к научной постановке задачи формирования рационального облика ИАС в целом. При этом рассмотрение ИАС обладает еще одним несомненным положительным качеством: применение данного подхода позволяет объективно формализовать как существующие области применения авиации (включая возможности их качественного изменения под влиянием технологического развития и самой авиации, и других областей техники), так и появление принципиально новых областей применения авиации.

Интегрированные авиационные системы могут быть выделены во многих отраслях (видах экономической деятельности) и сферах жизнедеятельности общества. Кроме рассмотренного примера санитарной авиации, формирование ИАС актуально для следующих случаев:

- пассажирские, грузовые и почтовые перевозки в отдаленных, труднодоступных и малонаселенных регионах (ОТДМР);
- агропромышленный комплекс (в том числе авиахимработы, точное земледелие);
- лесное хозяйство (в том числе авиалесоохрана, пожаротушение и т.п.);
- топливно-энергетический комплекс (ТЭК);
- внутригородской аэромобильный транспорт;
- мониторинг и инспектирование границ и территорий;
- дистанционное зондирование Земли, картографирование, разведка полезных ископаемых и биоресурсов;
- службы экстренного реагирования (обеспечение правопорядка);
- и др.

Во всех данных областях применения авиации актуален поиск рационального баланса между:

- с одной стороны, специфическими возможностями и особенностями применения авиационной техники;

## *Управление техническими системами и технологическими процессами*

– с другой стороны – технологическими потребностями различных отраслей и сфер жизнедеятельности, обусловленными их конечными целевыми задачами.

Собственно, это и наблюдается на протяжении всей истории авиации: изначально затраты на авиационную деятельность вообще отсутствовали, однако авиация, развиваясь, постепенно занимала все более значимое место в жизни общества. Именно такой взгляд на место авиации как средство решения некоторых конечных, фундаментальных проблем потребителей позволяет искать новые (перспективные) сферы применения авиации, корректно оценивать потенциальный спрос на авиационные работы и услуги, на продукцию авиастроения и смежных отраслей.

Что касается практического воплощения концепции ИАС, то она не исчерпывается реорганизацией на научной основе соответствующих областей жизнедеятельности. ИАС могут существовать как в форме сбалансированного комплекса разноплановых субъектов экономики (например, авиатранспортные предприятия и обслуживаемое население отдаленных населенных пунктов), так и в форме обособленного предприятия, консолидирующего всех участников отдельного вида деятельности (например, агропромышленное предприятие с авиационным подразделением). И если пространственная организация этих субъектов (а также, вероятно, сама технология решения целевых задач) будет меняться с учетом возможностей авиации, транспортных и других факторов, то именно их развитие в форме ИАС может оказать наибольшее благотворное влияние на социально-экономическое процветание общества в целом.

### **5. Математическая постановка задачи оптимального проектирования ИАС**

Состав ИАС, структура ее подсистем, характеристики их элементов должны быть целесообразными и эффективными с точки зрения целей применения и развития ИАС. Для достижения наилучших результатов, как уже было сказано, все перечисленные параметры ИАС следует оптимизировать совместно.

Рассмотрим более подробно математическую постановку задачи (1) применительно к задаче оптимизации ИАС, ее взаимосвязь и отличие от задачи оптимизации АС.

Пусть рассматривается  $K$  различных АС,  $1 \leq k \leq K$ , и для каждой  $k$ -й АС может быть сформулирована следующая задача оптимального проектирования:

$$(4) \quad \begin{cases} \min S_k(X_k), \\ F_k(X_k) = A_k, \\ G_k(X_k) \leq B_k; \end{cases}$$

где  $S_k(X_k)$  – векторная целевая функция затрат (по этой причине минимизируемая), рассчитываемая для  $k$ -й АС независимо от прочих АС;  $X_k = \{m_{1k}, \dots, m_{lk}, n_{1k}, \dots, n_{jk}, \bar{x}_{1k}, \dots, \bar{x}_{lk}, \bar{y}_{1k}, \dots, \bar{y}_{jk}\}$  – вектор проектных параметров, в состав которого входят управляющие переменные  $m_{1k}, \dots, m_{lk}$  как количества ВС  $l$  типов,  $1 \leq l \leq I$ , каждый из которых описывается вектором характеристик  $\bar{x}_{lk}$ , и  $n_{1k}, \dots, n_{jk}$  как количества  $J$  типов авиационной инфраструктуры,  $1 \leq j \leq J$ , каждый из которых описывается вектором характеристик  $\bar{y}_{jk}$ ;  $F_k(X_k)$  – векторная функция требуемых объемов работы  $A_k$  в  $k$ -й АС;  $G_k(X_k)$  – векторная функция требований (качественных ограничений)  $B_k$ , предъявляемых к  $k$ -й АС.

Тогда, в отличие от АС, постановка соответствующей оптимизационной задачи для «профильной» (отраслевой) ИАС будет иметь схожий вид, но со следующими изменениями:

$$(5) \quad \begin{cases} \min S_0(X_0), \\ F_0(X_0) = A_0, \\ G_0(X_0) \leq B_0; \end{cases}$$

где  $S_0(X_0)$  – векторная целевая функция затрат ИАС, в общем виде не совпадающая с любой из  $S_k$ ;  $X_0 = \{m_1, \dots, m_l, n_1, \dots, n_j, \bar{x}_1, \dots, \bar{x}_l, \bar{y}_1, \dots, \bar{y}_j, p_1, \dots, p_Q, \bar{z}_1, \dots, \bar{z}_Q\}$  – является вектором  $X_k$  тех же проектных параметров  $m_{1k}, \dots, m_{lk}, \bar{x}_{lk}, n_{1k}, \dots, n_{jk}$  и  $\bar{y}_{jk}$ , но расширенным за счет дополнительных переменных  $p_1, \dots, p_Q$  как количеств  $Q$  типов рассматриваемых объектов целевой отраслевой (неавиационной) инфраструктуры,  $1 \leq q \leq Q$ ,

## *Управление техническими системами и технологическими процессами*

каждый из которых также описывается вектором характеристик  $\vec{z}_q$ ;  $F_0(X_0)$  – векторная функция, определяющая требуемые объемы работы  $A_0$  в рассматриваемой ИАС, не совпадающая с любой из  $A_i$ ;  $G_0(X_0)$  – векторная функция ограничений  $B_0$ , предъявляемых к рассматриваемой ИАС, не совпадающая с любой из  $B_k$ .

Здесь необходимо подчеркнуть, что, в отличие от  $X_0$ , векторы  $A_0$  и  $B_0$  в рассматриваемой ИАС будут существенно другими по сравнению с соответствующими им  $A_k$  и  $B_k$ , в том числе не будут являться их расширенной суперпозицией.

Именно в этом и заключается особенность постановки задачи рационального формирования структуры и параметров ИАС. Для вектора  $A_0$  объем работы будет отличаться от объема авиационных работ, выполняемого непосредственно АС. Если вернуться к примеру медицинского обслуживания, то соответствующая ИАС оказывает услуги здравоохранения, в то время как АС санитарной авиации – только услуги медицинской транспортировки. Аналогичным образом в векторе  $B_0$  могут задаваться совершенно иные требования, чем в  $B_k$ : скажем, для ИАС предприятий трубопроводного транспорта целесообразно ставить условие прокачки определенных объемов нефти по заданному участку трубопроводной сети с минимальными потерями от утечек и аварий, а не условие выполнения патрульных облетов не меньше заданного объема, как для традиционной патрульной АС [8].

Вследствие этого векторы  $F_0$  и  $G_0$ , а также целевая функция  $S_0$  будут иметь вид, значительно отличающийся от  $F_k$ ,  $G_k$  и  $S_k$  (причем невозможно априорно выделить регулярные соотношения между ними).

### **6. Некоторые особенности формирования ИАС**

Исследование ИАС относится к задачам системного анализа. В соответствии с определениями этого раздела математической науки, ИАС следует классифицировать как очень сложную систему, т.е. систему, отличающуюся большой разветвленно-

стью связей и своеобразием отношений между элементами [1, 9]. Соответственно, проблематика, присущая ИАС, будет, при неизбежной специфичности, схожа с общей проблематикой системного анализа, как-то: формализация описания объектов проектирования, формирование различных проектных решений, методов их инженерного анализа и методов принятия решений, и т.д. Не останавливаясь подробно на данных аспектах, рассмотрим более детально только некоторые из них.

### *6.1. ЦЕЛЕПОЛАГАНИЕ РАЗВИТИЯ ИАС*

Как уже было отмечено выше, удовлетворение потребностей конечного заказчика авиационных работ или услуг как основного стейкхолдера является прямым стимулом для появления и развития новой авиационной техники. В то же время, помимо заказчика, существует еще несколько других групп стейкхолдеров, например:

- поставщики авиационных услуг или работ, если они выступают как независимые от заказчика субъекты;
- предприятия гражданской авиации: авиакомпании, аэропорты и другие предприятия, обеспечивающие ее функционирование;
- разработчики и изготовители авиационной техники, опелепродажных услуг и т.п., т.е. предприятия авиационной промышленности;
- государство (если оно не является заказчиком или поставщиком услуг) как выразитель интересов общества, в том числе третьих лиц, испытывающих на себе влияние авиационной деятельности: в сфере экологии, безопасности и т.п.;
- собственно указанные третьи лица.

В приведенном списке коммерческие предприятия преследуют, как правило, сугубо коммерческие цели. Внутри предприятий как сложных организмов также есть разные группы стейкхолдеров с различными интересами – владельцы (акционеры), наемные работники, менеджмент.

Что касается третьих лиц, то на них распространяются как положительные (улучшение транспортной доступности территорий), так и отрицательные внешние эффекты авиационной

## *Управление техническими системами и технологическими процессами*

деятельности (техногенные опасности, экологический ущерб и т.д.).

Государство (в идеале) балансирует интересы всех групп стейкхолдеров. И даже в тех секторах гражданской авиации, которые могут успешно развиваться на рыночной основе, оно, как регулятор, по крайней мере ограничивает рамки допустимого ущерба от авиационной деятельности, обеспечивает социальную защиту наемных работников всех отраслей и т.п.

Как результат, интересы всех перечисленных групп стейкхолдеров определяют целеполагание развития соответствующей ИАС. Поскольку этих групп несколько и их интересы не совпадают, целеполагание должно быть результатом компромисса между разнонаправленными интересами. Только такой системный взгляд на цели развития ИАС позволяет обеспечить корректное целеполагание, в том числе в части научно-технологического развития авиастроения.

### *6.2. ЦЕНТРАЛИЗАЦИЯ ПЛАНИРОВАНИЯ РАЗВИТИЯ ИАС*

Настоящая проблема вытекает из предыдущей относительно целеполагания развития ИАС и касается не только теоретической части вопроса, но и реализации ИАС на практике.

При формировании новой ИАС, строго говоря, неочевидно, какими характеристиками должны обладать ее элементы. Если на рынках магистральных пассажирских самолетов за десятилетия развития уже сложилась довольно стабильная структура, то в новых областях применения авиации, а также при значительном изменении традиционных областей, такой определенности еще нет. И на начальных этапах технологического развития определить оптимальные (и даже допустимые) характеристики элементов ИАС и связей между ними можно только централизованно [2]. С экономической точки зрения, появление независимых производителей инновационных продуктов/услуг (которые изначально рассматривались как соответствующие элементы ИАС) становится целесообразным, только когда ожидаемый объем реализации этих продуктов/услуг превысит определенный порог.

Данное обстоятельство может значительно повысить риск ошибки при создании ИАС и потребовать более совершенных методов моделирования. В частности, может потребоваться разработка не статической, а динамической математической модели ИАС. В ней заказчику авиационных работ или услуг, взяв на себя роль системного интегратора, на первом этапе придется самостоятельно обеспечивать себя необходимыми компонентами и услугами (недостающими элементами системы), и лишь по мере созревания рынка передавать эти функции на аутсорсинг.

С практической же точки зрения, из приведенных рассуждений вытекает необходимость изначально предлагать клиентам комплексные решения в виде интегрированных систем, а не искать потребителей отдельно взятых ЛА. Например, производитель аэротакси для внедрения своих ЛА будет вынужден принять непосредственное участие в создании (по крайней мере организовать создание по заказу городских властей или частного инвестора) городской авиатранспортной системы со всей необходимой управляющей, сервисной и логистической инфраструктурой, нормативной и правовой базой. Практическая реализация ИАС будет означать создание в городе или агломерации системы «под ключ», представляющей собой комплексное решение для пользователей транспортных услуг. Лишь потом, когда подобные системы будут сформированы, а ЛА станут их стандартными элементами с обоснованными и проверенными на практике требованиями к значениям их характеристик, появятся отдельные рынки данных классов ЛА, и возникнут условия для развития конкурентных рынков этих изделий.

## **7. Применение методологии ИАС на примере ТЭК**

В качестве примера применения предлагаемого методического аппарата может быть рассмотрена ИАС ТЭК (обслуживание магистралей нефте-/газопровода).

В состав такой ИАС входят малые патрульные беспилотные летательные аппараты (БПЛА), специализированная наземная инфраструктура базирования БПЛА, дежурные бригады оперативного реагирования, каждая из которых оснащена одним спе-



*Управление техническими системами  
и технологическими процессами*

циализированным 3÷4-местным вертолетом (винтокрылым летательным аппаратом, ВКЛА).

Вдоль магистрали длиной  $L_{\text{маг}}$  патрулируют малые БПЛА, челночно перемещаясь от одного пункта базирования до другого и обратно. Пункты базирования удалены друг от друга на расстояние  $L_{\text{БПЛА}}$ , и условно считается, что это расстояние меньше дальности полета БПЛА (т.е. ограничение на дальность полета БПЛА не ставилось). Общее количество БПЛА и пунктов базирования –  $N_{\text{БПЛА}}$ . БПЛА характеризуются крейсерской скоростью  $V_{\text{БПЛА}}$  и стоимостью летного часа  $c_{\text{л.ч.}}$ . Годовая стоимость содержания одного пункта базирования –  $c_{\text{пл.}}$ .

При обнаружении утечки нефти БПЛА в автоматическом режиме вызывает ремонтную бригаду. Пункты базирования ремонтных бригад также расположены вдоль магистрали, и расстояние между ними определяется радиусом действия ВКЛА.

Оператор нефтегазовой магистрали испытывает потребность не в патрулировании как таковом, а в минимизации суммарных затрат на патрулирование и потерь от утечек нефти. Эффект от внедрения более эффективных ВС в сфере патрулирования магистралей будет наибольшим, если типаж, численность и налет парка летательных аппаратов (ЛА), а также облик потребной инфраструктуры будут способствовать скорейшему обнаружению аварийных ситуаций и минимизации ущерба от них. Функция годовых затрат  $S$  оператора магистрали в этом случае будет иметь вид:

$$(6) \quad S = C_{\text{ВКЛА}} + C_{\text{БПЛА}} + C_{\text{пл}} + C_{\text{уб}} + C_{\text{ш}},$$

где  $C_{\text{ВКЛА}}$  – стоимость годовых затрат на содержание ремонтных бригад, включая эксплуатацию ВКЛА;  $C_{\text{БПЛА}}$  – стоимость годового налета БПЛА;  $C_{\text{пл}}$  – годовая стоимость содержания пунктов базирования БПЛА;  $C_{\text{уб}}$  – стоимость вытекшей нефти как потерянного продукта;  $C_{\text{ш}}$  – экологический штраф.

Вводится предположение, что повреждения трубопровода могут происходить равновероятно в любом месте. В этом случае затраты  $C_{\text{ВКЛА}}$  зависят только от длины трубопровода и количества аварийных ситуаций и могут быть рассчитаны априорно независимо от других статей затрат. Иначе говоря, для заданно-

го трубопровода затраты  $C_{\text{ВКЛА}}$  входят в выражение для  $S$  как константа, и по этой причине  $C_{\text{ВКЛА}}$  далее подробно не рассматриваются.

Пусть  $T_0$  – регулярность (дни), с которой БПЛА совершают облет магистрали. Тогда

$$(7) \quad C_{\text{БПЛА}} = C_{\text{л.ч.}} \times \frac{L_{\text{маг}}}{V_{\text{БПЛА}}} \times \frac{365}{T_0}; \quad C_{\text{ш}} = N_{\text{БПЛА}} \times c_{\text{ш}};$$

$$C_{\text{уб}} = c_{\text{н}} \times m_{\text{н}}^t \times t_0; \quad C_{\text{уб}} = c_{\text{ш}} \times (m_{\text{н}}^t \times t_0)^2.$$

Здесь  $c_{\text{н}}$  – стоимость единицы нефти;  $m_{\text{н}}^t$  – количество нефти, в среднем вытекающей из места повреждения за единицу времени;  $t_0$  – время, прошедшее с момента возникновения утечки до ее устранения;  $c_{\text{ш}}$  – коэффициент экологического штрафа за единицу вытекшей нефти. При этом общий штраф  $C_{\text{уб}}$  начисляется пропорционально квадрату количества вытекшей нефти: считается, что нефть растекается по окружающей территории как по площади, тем самым размер ущерба окружающей среде нарастает весьма быстрыми темпами:

$$(8) \quad m_{\text{н}}^t = w \times \frac{L_{\text{маг}}}{1000} \times \rho,$$

где  $w$  – вероятность повреждения трубопровода в год из расчета на 1000 км его длины;  $\rho$  – осредненный поток вытекающей нефти.

Варьируемыми параметрами ИАС являются регулярность облета  $T_0$  и количество БПЛА  $N_{\text{БПЛА}}$ . Отсюда  $t_0$  вычисляется как среднее время обнаружения в зависимости от того, в каком случае обнаружение утечки произойдет раньше:

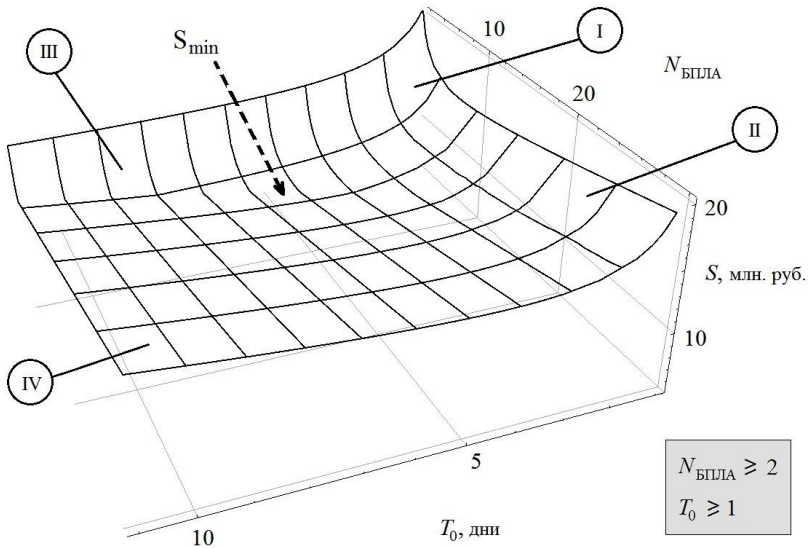
$$(9) \quad t_0 = t_p + \frac{1}{2} \times \begin{cases} T_0, & \frac{L_{\text{маг}}}{N_{\text{БПЛА}} \times V_{\text{БПЛА}}} \leq T_0, \\ \frac{L_{\text{маг}}}{N_{\text{БПЛА}} \times V_{\text{БПЛА}}}, & \frac{L_{\text{маг}}}{N_{\text{БПЛА}} \times V_{\text{БПЛА}}} > T_0. \end{cases}$$

Здесь  $t_p$  – средняя поправка, включающая время прибытия ремонтной бригады и остановки утечки, невозможность обнаружения утечки в ночное время и т.д.

*Управление техническими системами  
и технологическими процессами*

В качестве примера при расчете затрат использовались следующие значения параметров:  $L_{\text{маг}} = 2400$  км,  $V_{\text{БПЛА}} = 60$  км/ч,  $c_{\text{л.ч.}} = 500$  руб./л.ч.,  $C_{\text{пл}} = 240$  тыс. руб./год,  $c_{\text{н}} = 25$  тыс. руб./м<sup>3</sup>,  $w = 0,08$ ,  $\rho = 0,0044$  м<sup>3</sup>/с (или 380 м<sup>3</sup>/день, нефиксируемый приборами расход нефти),  $c_{\text{ш}} = 2500$  (задан как 10% от стоимости вытекшей нефти за день),  $t_{\text{р}} = 12$  часов.

Общий вид поверхности  $S$  имеет вид (рис. 4, целочисленные величины трактуются как непрерывные).



*Рис. 4. Затраты ИАС в зависимости от регулярности облета  
магистралей и количества патрульных БПЛА  
(не включая затраты на ВКЛА)*

Приведенные случаи соответствуют:

- I – «малое количество БПЛА, высокая регулярность облета»;
- II – «большое количество БПЛА, высокая регулярность облета»;
- III – «малое количество БПЛА, низкая регулярность облета»;
- IV – «большое количество БПЛА, низкая регулярность облета».

Во всех приведенных случаях рост затрат связан или с ростом расходов на авиационную составляющую ИАС, или с ро-

стом убытков вследствие потерь транспортируемой нефти. Минимум затрат  $S_{min} \sim 9,2$  млн руб./год – это область поверхности, которая соответствует параметрам  $T_0 = 6 \div 7$  дней,  $N_{БПЛА} = 6$ .

Таким образом, для заданного примера выбор оптимальных параметров ИАС может привести к экономии на уровне 20÷50%:  $\sim 10$  млн руб./год вместо 12÷15 млн руб./год при неоптимальном выборе (области I÷IV на рис. 4).

## 8. Выводы

На современном этапе развития авиации при планировании или прогнозировании ее развития становятся устаревшими расчетные методы, в которых оперируют изолированными воздушными судами, авиационными комплексами или даже авиационными системами, выполняющими заданные объемы авиационных работ и услуг. Наиболее комплексным подходом является рассмотрение интегрированных авиационных систем (ИАС), в рамках которых возможна совместная оптимизация состава и численности парка авиационной техники, необходимой авиационной инфраструктуры и неавиационных объектов целевого заказчика авиационных работ или услуг.

В основе подхода ИАС лежит тезис о решающем значении целеполагания применения авиации с точки зрения конечных заказчиков. И поскольку любая ИАС формируется исходя из их потребностей, уровень ИАС в иерархии уровней оптимизации ЛА располагается выше, чем уровень просто авиационных систем, что, с одной стороны, повышает сложность решаемой задачи, но, с другой стороны, позволяет найти более рациональные в глобальном плане решения.

При этом обсуждаемые проблемы значимы не только для авиации. Сама идея ИАС заключается в совместной оптимизации авиационной и неавиационной составляющих отрасли «клиента». Развитие интегрированных авиационных систем как рациональное распределение задач между ЛА различных типов и рациональное размещение необходимой авиационной и неавиационной инфраструктуры может оказать положительное влияние на пространственную организацию общества, экономи-

## *Управление техническими системами и технологическими процессами*

ки, государства. Современные средства математического моделирования позволяют найти необходимые решения в этой сфере быстрее, чем методом проб и ошибок, который при этом оказывается еще и неприемлемо дорогостоящим и рискованным.

Развитие методологии ИАС позволит адекватно оценивать истинные потребности заказчиков авиационных работ и услуг, возможности их удовлетворения посредством авиации, прогнозировать потенциальные объемы применения авиации, в том числе находить новые перспективные рынки реализации как авиационных работ и услуг, так и продукции авиастроения. Необходимость централизованно формировать и исследовать интегрированные авиационные системы становится очевидной именно при переходе к новым технологическим укладам – как в авиации, так и в областях ее целевого применения.

### **Литература**

1. АНТОНОВ А.В. *Системный анализ* – М.: Высшая школа, 2004. – 454 с.
2. ДУТОВ А.В., КЛОЧКОВ В.В. *Стратегическое управление развитием авиационных технологий: проблемы и современные решения* // Экономический анализ: теория и практика. – 2013. – №48(351). – С. 2–15.
3. ЖЕЛТОВ С.Ю., КИСЛИЦЫН Ю.Д., САМОЙЛОВ Д.В., ХОХЛОВ С.В. *Современные подходы в моделировании авиационных систем* // Сборник тезисов докладов IV Всероссийской научно-технической конференции. – М.: ГосНИИ АС. – 2020. – С. 5.
4. ЖЕРЕБИН А.М., КРОПОВА В.В., МАЛАФЕЕВ И.В. *Концептуальные основы и модельно-методическое обеспечение процесса управления созданием научно-технического задела с позиции формирования рационального состава парка самолетов заданного функционального назначения* // Сборник тезисов докладов IV Всероссийской научно-технической конференции. – М.: ГосНИИ АС. – 2020. – С. 12-13.

5. ЕГЕР С.М., ЛИСЕЙЦЕВ Н.К., САМОЙЛОВИЧ О.С. *Основы автоматизированного проектирования самолетов*. – М.: Машиностроение, 1986. – 232 с.
6. ИСАЕВ А.С., СТАВРОВСКИЙ Б.И. *Основы эффективного анализа и синтеза авиационных комплексов: Учебное пособие*. – М.: МАИ, 1983. – 78 с.
7. КЛОЧКОВ В.В., ЕГОШИН С.Ф. *Задачи развития санитарной авиации и совершенствования пространственной организации здравоохранения в России // Россия: тенденции и перспективы развития. Ежегодник*. – 2020. – Вып. 15, ч. 2. – С. 628–637.
8. КЛОЧКОВ В.В., НИКИТОВА А.К. *Методы прогнозирования спроса на беспилотные летательные аппараты и работы по воздушному патрулированию // Проблемы прогнозирования*. – 2007. – №6. – С. 144–151.
9. НОВИКОВ Д.А. *Теория управления организационными системами*. – М.: МПСИ, 2005. – 584 с.
10. СЛИВИЦКИЙ А.Б. *Комплексный анализ целеполагания в системе стратегического планирования в области авиационной деятельности // Россия: тенденции и перспективы развития. Ежегодник / Под ред. В.И. Герасимова*. – Институт научной информации по общественным наукам РАН, отдел научного сотрудничества, 2019. – С. 475–482.
11. ТОПОРОВ Н.Б. *Применение моделирования для внешнего проектирования сложных организационно-технических авиационных систем // Труды Международной научной конференции СРТ2019, 13-17 мая 2019 г., Царь-Град, Московская область, Россия*. – 2019. – С. 50–56.
12. ТОПОРОВ Б.П., ГОРЛОВ В.М. *Проблемы формирования рационального облика и типажа сложных технических систем, функционирующих в условиях конфликта и неопределённостей // Труды IV Международной конференции «Проблемы управления и моделирования в сложных системах», Самара, Россия*. – 2002.

13. ADJEKUM D.K., TOUS M.F. *Assessing the relationship between organizational management factors and a resilient safety culture in a collegiate aviation program with Safety Management Systems (SMS)* // Safety science. – 2020. – Vol. 131. – DOI: 10.1016/j.ssci.2020.104909.
14. DRAY L.M., KRAMMER P., DOYME KH., WANG B., AL ZAYAT K., O'SULLIVAN A., SCHAFER A.W. *AIM2015: Validation and initial results from an open-source aviation systems model* // Transport policy. – 2019. – Vol. 79. – P. 93–102.
15. EMMANOUIL K. *Reliability in the era of electrification in aviation: A system approach* // Microelectronics reliability. – 2020. – Vol. 114.
16. FEI X., BIN CH., SIMING ZH. *A methodology of requirements validation for aviation system development.* – August 2020. – P. 4484–4489. – DOI: 10.1109/ccdc49329.2020.9164301.
17. FUJIMOTO R.M. *Parallel and distributed simulation systems.* – Wiley, 2000. – 320 p.
18. PRITCHETT D. *Base: an acid alternative* // ACM Queue. – N.Y.: ACM, 2008. – Vol.6, No.3. – P. 48–55.
19. SAVKIN A.V., HUANG H. *Asymptotically optimal deployment of drones for surveillance and monitoring* // Sensors. – 2019. – No. 19. – 11 p.
20. WITTMER A., BIEGER T. *Fundamentals and structure of aviation systems* // Aviation systems. Management of the integrated aviation value chain. – Hardcover, 2011. – Vol. XVII. – P. 5–38.

## **INTEGRATED AVIATION SYSTEMS**

**Vladislav Klochkov**, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Doctor of Science (klochkovvv@nrczh.ru).

**Nikolay Toporov**, National Research Center “Zhukovskiy Institute”, Zhukovskiy, Doctor of Science, professor (toporov@nrczh.ru).

**Sergey Yegoshin**, National Research Center “Zhukovskiy Institute”, Zhukovskiy, (sergey4791@yandex.ru).

*Abstract: Improving the efficiency of management of aircraft industry scientific and technological development at this stage is possible through the transition to the consideration of large-scale systems with a large number of "degrees of freedom".*

*For the first time methodological approach to forecasting and strategic planning of scientific and technological development of aircraft industry, based on analysis and synthesis of integrated systems, rather than individual aircraft or aviation systems, has been proposed and developed. These systems are designed to address high-level objectives and include fleets of various aircraft, as well as infrastructure and control subsystems, including non-aviation subsystems of the target customer. The research methodology is based on the systems theory, system analysis and synthesis methods, and mathematical modeling. The proposed approach allows us to reveal the potential fields of application of prospective aviation in transport and various sectors of the national economy, as well as estimate the potential scale of aviation. The results of modeling and optimization of integrated aviation systems should form the basis for planning of priority research and development in aircraft industry, including the formation of state programs of the Russian Federation, the development strategy of scientific and industrial organizations of aircraft industry.*

Keywords: integrated aircraft system, modelling, optimization, aircraft design, systems analysis, scientific and technological development of aircraft industry.

УДК 051.7 + 629.1

ББК 39.52

DOI: 10.25728/ubs.2021.90.5

*Статья представлена к публикации  
членом редакционной коллегии Р.М. Нижегородцевым.*

*Поступила в редакцию 11.01.2021.*

*Опубликована 31.03.2021.*



## **ПРОЕКТИРОВАНИЕ СЕТЕЙ СВЯЗИ БЫСТРОГО РАЗВЕРТЫВАНИЯ С ИСПОЛЬЗОВАНИЕМ ПРОГРАММНОГО ПРИЛОЖЕНИЯ АНАЛИЗА КОНФИГУРАЦИИ СЕТИ**

**Блинова О. В.<sup>1</sup>, Васьковский С. В.<sup>2</sup>, Рыков Я. В.<sup>3</sup>**  
(ФГБУН Институт проблем управления  
им. В.А. Трапезникова РАН, Москва)

*Предложен новый подход к проектированию беспроводных сетей связи быстрого развертывания при невозможности полного покрытия территории сетью в заданных условиях, суть которого заключается в определении оптимальных мест расположения комплектов сетевого оборудования. При поиске целевой функции используется суммарный по всем абонентам уровень доступа к сети в зависимости от времени, к которому возможно добавлять различные коэффициенты для учета особенностей конкретной задачи. Для практического применения рассматриваемого подхода создан программный продукт, реализующий разработанную процедуру поиска с применением имитационного моделирования и генетических алгоритмов. Рассмотрен алгоритм работы приложения и приведены примеры результатов моделирования. Задача размещения сетевых устройств в заданных условиях решена впервые, созданный программный продукт не имеет аналогов. Полученные результаты не сводятся к однозначному решению, но применимы для анализа и исследования задачи быстрого проектирования беспроводных сетей связи в областях, где нет покрытия мобильной сетью связи, для обеспечения связью спасательных работ или туристических маршрутов в горной местности.*

Ключевые слова: беспроводные сети связи, сети связи с неполным покрытием, обеспечение связью спасательных работ, сети связи быстрого развертывания.

### **1. Введение**

Сети связи быстрого развертывания применяются для временного обеспечения связью (мобильной или доступом к сети Интернет) пользователей, передвигающихся по территории с отсутствующим или неполным покрытием сетью связи. Такие

---

<sup>1</sup> Ольга Викторовна Блинова, н.с. (blinova\_olga\_v@mail.ru).

<sup>2</sup> Сергей Владимирович Васьковский, к.т.н., с.н.с. (v63v@yandex.ru).

<sup>3</sup> Ярослав Владимирович Рыков, вед. инж. по АСУП (yarykov@gmail.com).

сети могут применяться для обеспечения связью зон ликвидации последствий чрезвычайных происшествий, обеспечения связью туристических и спортивных мероприятий. Решение задачи проектирования беспроводной сети в подобных ситуациях часто осложнено дефицитом времени и возможностей для проектирования, ограниченным набором оборудования, большой протяженностью маршрутов. Авторы рассматривают задачи, в которых обеспечение полного покрытия территории сетью не представляется возможным, но обеспечение эпизодического доступа абонентов к сети является целесообразным. Обеспечение на маршрутах движения определенных зон доступа позволит абонентам передавать необходимые данные и получать рекомендации или указания. Данная статья является продолжением и развитием результатов, полученных в работах [2, 3, 4].

При необходимости создания сети для подключения подвижных абонентов, тем более сетей быстрого развертывания, выбор однозначно осуществляется среди беспроводных систем. Территория, на которой необходимо развернуть сеть, может отличаться сложными условиями – отсутствием электропитания, сложной климатической обстановкой, отсутствием дорог нужного качества. Оптимальным решением может быть использование специально разрабатываемых для подобных ситуаций готовых программно-аппаратных комплексов, включающих независимые источники электропитания и сетевое оборудование. Примерами таких комплексов могут быть системы связи и безопасности особого назначения МИКРАН [10], мобильные станции спутниковой связи [6], мобильный комплекс диспетчеризации радиосвязи и сопряжения с другими каналами коммутации МИРАН® «Кейс-Ч» [19] и другие подобные решения [1, 15, 18].

В рассматриваемых условиях существует проблема правильного размещения сетевого оборудования на местности для максимально эффективного его использования. В ряде ситуаций необходимость создания сети не задана заранее или нет возможности заранее изучить местность и провести расчеты. Возникает нетривиальная задача проектирования сети связи, решение которой должно быть получено в течении короткого промежутка времени. Для быстрого нахождения решения определен-

ного класса задач целесообразно использование программно-инструментальных средств.

Существуют программные продукты, позволяющие упростить и ускорить проектирование беспроводных сетей для офисов и помещений. Обзор таких программных решений есть в [12, 13]. Примерами подобных моделей являются [14, 16, 17]. Обычно пользовательские задачи требуют бесперебойной и высокоскоростной связи, и требования к сети формулируются как полное покрытие заданной области зонами доступа, распределение нагрузки между ними, обеспечение максимальной помехозащищенности, учет материалов стен, перекрытий и особенностей распространения сигнала. Как правило, такие сети проектируются стационарно и рассчитаны на несколько лет работы. Эта категория программных продуктов не может быть эффективно применена для разработки сетей связи с неполным покрытием, так как в рамках существующих инструментов невозможно учесть многие критически важные факторы, например, траектории движения отдельных абонентов.

Среди задач с неполным покрытием территории сетями связи можно выделить класс задач по обеспечению связи вдоль транспортных магистралей. Обзор работ по этой теме приведен в статье [4], примерами работ с обеспечением связью линейной траектории являются [5, 7, 9, 11].

Подобные задачи возникают при обеспечении связью речного или морского транспорта [8]

Перечисленные задачи обладают рядом общих черт с задачами построения сетей быстрого развертывания: применение беспроводных технологий, подвижность абонентов, проблемы размещения точек доступа, привязка к траектории движения. Но в рассматриваемом авторами классе задач дополнительно учтен ряд значимых факторов, а именно:

- сложность условий функционирования оборудования и высокие требования к его надежности;
- относительно небольшое число пользователей и невысокая их плотность;
- трудности обеспечения связи между устройствами – ограниченная возможность использования проводной связи, не-

устойчивое покрытие сети, наличие помех, большие расстояния между устройствами;

- сложность формализации задачи и т.д.

Из вышесказанного следует, что решение рассматриваемой авторами задачи не может быть сведено к использованию существующих инструментов, и работы над разработкой новых подходов являются актуальными. Возможный способ решения проблемы проектирования сетей быстрого развертывания с неполным покрытием территории – разработка программного продукта, позволяющего частично автоматизировать процесс решения задачи о размещении на территории сетевого оборудования.

## 2. Постановка задачи

При формулировке задачи и построении модели за основу была взята ситуация перемещения туристических групп по маршрутам в горной местности (местности со сложным рельефом) или спасателей при обнаружении чрезвычайного происшествия. Такой выбор был сделан в связи с реально существующими задачами обеспечения связи при проведении различных мероприятий в малонаселенной гористой местности. Горные вершины как препятствия для размещения устройств и техники могут быть легко замещены другими препятствиями, привязанными к конкретному району, рельефу, условиям. Например, это может быть болотистая местность.

Задача формулируется следующим образом:

Рельеф местности задан в виде поверхности в декартовых координатах размерами  $x_0$  на  $y_0$  на  $z_0$ . На ортогональной проекции поверхности на горизонтальную плоскость задаются  $k$  маршрутов абонентов в виде отрезков, заданных двумя точками, для каждого абонента задана постоянная относительно плоскости скорость движения  $v_j$ ,  $j = 1, \dots, k$ . Абоненты движутся по маршрутам циклично. Также даны  $n$  комплектов сетевого оборудования с радиусом действия  $R_i$ ,  $i = 1, \dots, n$ . Необходимо найти координаты  $(x_i, y_i)$  размещения комплектов сетевого оборудования, при которых обеспечивался бы максимальный уровень доступности сети для абонентов. Уровень доступа  $j$ -го абонента к  $i$ -му узлу в момент времени  $t$  задается следующим образом:

$$(1) p_{ij} = e^{-d/2R_i^2}, \quad d = (x_i - x_j)^2 + (y_i - y_j)^2,$$

где  $x_j$  и  $y_j$  – координаты  $j$ -го узла в момент времени  $t$ .

Выбор целевой функции в данных условиях является не тривиальной задачей. Одним из возможных решений является использование средней вероятности доступа к сети абонентов за некоторый интервал времени, что требует строгости определения и подразумевает невозможность использования каких-либо весовых коэффициентов. Важно отметить, что рассчитываемые значения уровня доступа являются очень приближительными, так как в предлагаемой модели учитываются только расстояния между сетевыми устройствами и абонентами, но нет учета распространения сигнала по сложному рельефу, нет возможности учитывать погодные условия, совместимость оборудования и прочие факторы, влияющие на распространение сигнала и обеспечение связи. В этой связи в качестве целевой функции был выбран условный параметр  $w$ , названный «успешностью» конфигурации. В приведенных результатах моделирования он рассчитан следующим образом:

$$(2) w = \sum_{i=1}^n \sum_{j=1}^k p_{ij},$$

где  $n$  – число точек доступа;  $k$  – число абонентов;  $p_{ij}$  – вероятность доступа  $j$ -го абонента к  $i$ -й точке доступа.

Для удобства использования полученное значение нормируется следующим образом:

$$(3) W = 5w / (1 + 0,02w).$$

Этот показатель не имеет физического смысла; предполагается, что метод его расчета может выбираться индивидуально в зависимости от задачи. Например, при приоритетной необходимости обеспечить связь конкретному абоненту его вклад в общий показатель «успешности» может быть умножен на некоторый коэффициент, характеризующий уровень приоритета. Если оказывается, что некоторая точка доступа не сможет на практике обеспечить заявленные характеристики, для всех абонентов полученные расчетные уровни доступа к этой точке можно снизить на определенный коэффициент. Таким образом, подстраиваясь под различные индивидуальные условия, пользователь может адаптировать полученные результаты для конкретной задачи с помощью простых коэффициентов. Безусловно,

но, при такой постановке возникает вопрос о значимости получаемых расчетных характеристик. Но авторы ставят перед собой задачу разработки программного продукта, являющегося помощником экспертам на местности, инструментом для экспериментов, позволяющим проследить взаимозависимости разных факторов.

Тестирование модели с использованием данной целевой функции привело к выявлению следующего недостатка: эффективным получается постановление всех сетевых устройств в одном месте, что явно противоречит практической значимости. Для устранения отмеченного недостатка в модель были добавлены ограничение на размещение устройств выше определенного уровня и функция отталкивания сетевых устройств друг от друга, рассчитываемая на каждом шаге моделирования по формуле

$$(4) \quad b = 0,4e^{-0.004l_j} \left( \frac{100}{1+0,25l_j} \right).$$

Значения ограничений подобраны эмпирически, они не учитываются при расчете значений целевой функции, но влияют на совершаемое перемещение сетевого устройства.

Целью моделирования является максимизация значения целевой функции  $W$ . Ее значение рассчитывается на каждом шаге моделирования, на основе полученных значений выполняется шаг (смещение сетевых устройств), процесс продолжается до остановки моделирования. В процессе моделирования точки доступа смещаются непрерывно, решение о использовании подвижных точек доступа и построении маршрута их движения или о выборе стационарного места размещения принимает эксперт на основе полученных результатов.

### **3. Программная реализация**

Для решения задачи был создан программный продукт, решающий сформулированную задачу с использованием генетических алгоритмов и имитационного моделирования. При выборе этого метода были учтены следующие факторы:

1. Нет необходимости находить абсолютный максимум целевой функции. Для практических целей вполне достаточно решения, близкого к оптимальному. Использование генетических

алгоритмов позволяет плавно приближаться к максимуму и остановиться в поиске в любой момент, приняв полученное промежуточное решение за конечный результат.

2. Для выбора новых точек размещения оборудования генетические алгоритмы одновременно используют как вероятностные, так и детерминированные правила, что дает значительно больший эффект, чем каждый из этих методов в отдельности.

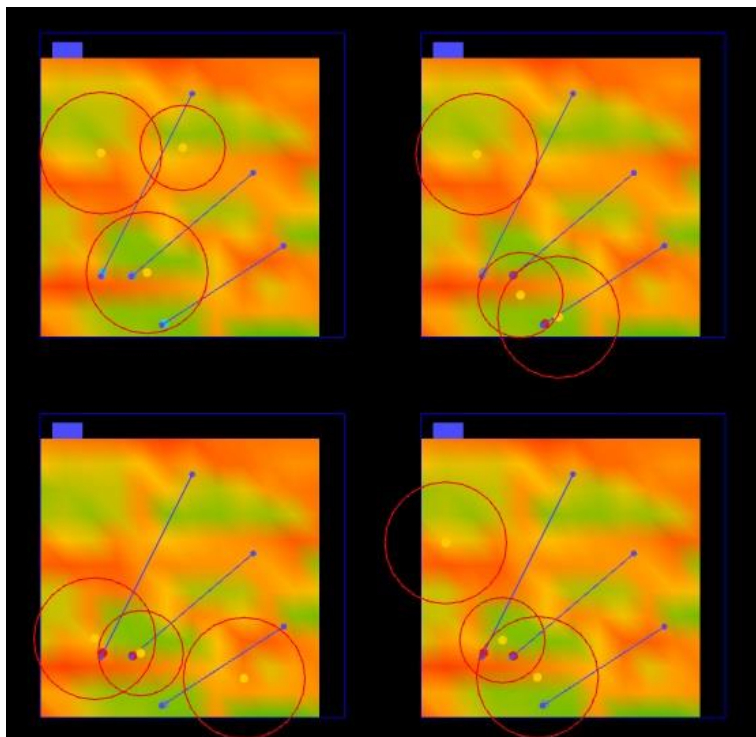
3. Генетические алгоритмы отличаются низкой вычислительной мощностью и при грамотной формулировке целевой функций и базовых операций позволяют добиться хорошего решения за короткое время.

4. Особенности поставленной задачи хорошо формулируются в терминологии генетических алгоритмов.

Для ввода исходных данных и управляющих параметров сформирован конфигурационный файл. В нем маршруты определены в виде отрезков, заданных двумя точками (ломаные линии можно задавать несколькими отрезками). Там же указывается число и радиус действия доступных сетевых устройств, а также вероятности кроссинговера, мутации. В рассматриваемом варианте не выполнена привязка рельефа местности к реальным географическим картам.

Для решения поставленных задач и исследования местности на пригодность размещения сетевого оборудования был создан алгоритм создания случайного горного рельефа. Он генерируется при помощи тригонометрических функций, создающих колебания по оси  $x$  и  $y$ . Для получения правдоподобного рельефа используются три функции, создающие выпуклости различного масштаба, и их сумма определяет генерируемый рельеф. Для наглядности высотность местности обозначена цветовым градиентом от зеленого (низины) к темно-красному цвету (горы) (рис. 1). При запуске программы генерируется 8 случайных вариантов размещения оборудования. Четыре лучших из них по значению целевой функции выводятся на экран.

Для наглядности и удобства изучения предлагаемых конфигураций реализовано трехмерное представление с возможностью поворачивать изображение и изменять угол наклона, также возможно увеличить одну выбранную конфигурацию (рис. 2).

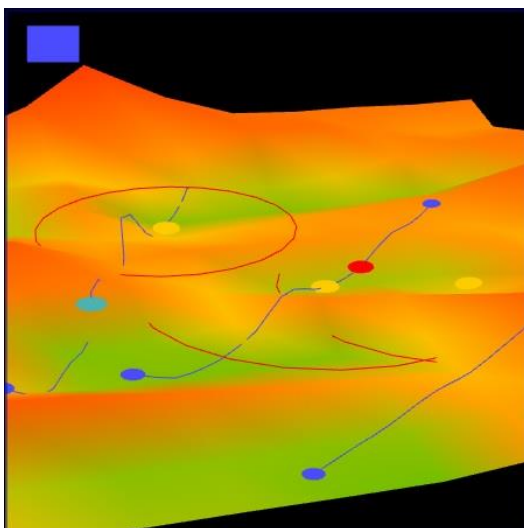


*Рис. 1. Начальный вариант размещения оборудования, визуализация исходных данных*

Для использования генетических алгоритмов необходимо определить физический смысл базовых понятий: популяции, особи, хромосомы, а также смысл процедур кроссовера и мутации. Популяцией является множество сгенерированных вариантов размещения сетевого оборудования, особью – один такой вариант, хромосомой – множество координат размещения сетевых устройств. Для процедуры кроссовера выбираются две удачные конфигурации и на основе их координат размещения точек доступа выбираются новые координаты, средние между координатами родителей. Для процедуры мутации одна из координат одной из точек доступа меняется на случайную величину (точка доступа смещается на случайное расстояние в случайную



сторону). Возникновение мутаций позволяет получить большее разнообразие конфигураций. Для формирования нового поколения рассматриваются все конфигурации (и родительские, и дочерние), и лучшие из них по значению целевой функции попадают в следующую популяцию, размер популяции остается неизменным.



*Рис. 2. Сгенерированная конфигурация  
в трёхмерном представлении*

В связи с непрерывным движением абонентов значения целевой функции не могут непрерывно увеличиваться, и их изменения связаны с точкой нахождения абонента на маршруте. Это может привести к ошибкам при оценке качества той или иной конфигурации, поэтому для дополнительной иллюстративности в систему добавлен показатель  $W_0$ , который учитывает текущее значение целевой функции и предыдущие значения (до 80-го) с убывающими весами.

После запуска программы и генерации начальных вариантов размещения оборудования для каждой из точек доступа определяется вектор возрастания целевой функции (приближе-

ния к абонентам). Точки доступа перемещаются по найденному вектору небольшими случайными по модулю шагами. Таким образом каждая из конфигураций стремится к возможному максимуму. Значения вектора отталкивания рассчитываются для каждого из сетевых устройств по отношению ко всем остальным, найденные значения суммируются, результирующий вектор складывается с шагом перемещения согласно целевой функции, проверяется соответствие ограничению по высоте и после этого выполняется перемещение сетевого устройства. Если планируемый шаг приведет к нарушению ограничения, сетевое устройство не перемещается.

Для оценки и анализа полученных конфигураций результаты моделирования заносятся в таблицы данных. Сводная таблица содержит номер записи, время в секундах от начала моделирования и сведения о четырех лучших конфигурациях – их id и значение целевой функций, а также функции  $W_0$ , обладающей «инерцией» и позволяющей лучше оценить динамику изменений. Также в системе есть возможность отображения последней операции, связанной с конфигурацией (начальная или возникшая в результате мутации или кроссинговера). Сортировка конфигураций осуществляется раз в 5 секунд, поэтому иногда у первой конфигурации значение целевой функции не максимально. Форма результирующей таблицы для приведенной выше конфигурации приведена в таблице 1.

Табл. 1. Форма сводной таблицы по лучшим конфигурациям

n	Time	1st place			2nd place			3rd place			4th place		
		id	w	w0	id	w	w0	id	w	w0	id	w	w0
1	0,00	0	0,17	nan	1	0,12	nan	2	0,09	nan	3	0,20	nan
2	1,06	0	0,20	nan	1	0,20	nan	2	0,11	nan	3	0,24	nan

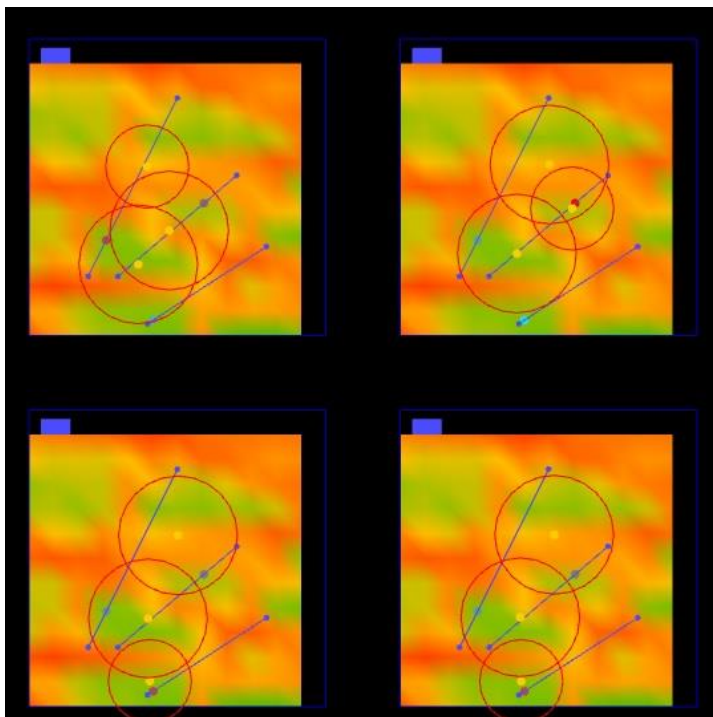
Результирующей таблицы недостаточно для получения подробных сведений об уровне доступа, который реально обеспечивается абоненту, и о координатах расположения сетевых устройств. Для просмотра этих сведений формируется таблица для каждой из четырёх лучших конфигураций. Все сформированные таблицы совместимы с продуктами Microsoft Office

и могут быть перенесены в эти или другие программные продукты для дальнейшей обработки, например, построения графиков уровня доступа в зависимости от расположения. Таблица, описывающая конкретную конфигурацию, содержит номер записи, время моделирования, координаты всех абонентов, координаты всех сетевых устройств, уровень доступа, обеспечиваемый каждому абоненту, и расстояние до ближайшей к нему точки доступа. Форма этой таблицы представлена ниже (таблица 2.) Таблицы 1 и 2 совместимы по номеру записи и времени моделирования и при необходимости их можно объединить. Время моделирования указано программное, но оно синхронизируется с реальным временем движения абонентов по их скорости. По данным этой таблицы легко видеть, какой уровень доступа обеспечивается каждому из абонентов в конкретный момент времени.

*Таблица 2. Сведения о конфигурации*

	People Coords			Node Coords			Signal Power			Dist To Nearest Node			
Time	Person 0	Person 1	Person 2	Node 0	Node 1	Node 2	Pers 0	Pers 1	Pers 2	Pers 0	Pers 1	Pers 2	
00,00	1500; 1000;	2000; 200;	1000; 1000;	1213; 139;	2325; 641;	1584; 2941;		1,351	1,593	1,079	901	548	887
11,06	1569; 1059;	2131; 285;	1116; 1233;	1287; 701;	1830; 799;	1533; 2891;		1,868	1,481	1,611	368	595	559
22,04	1634; 1113;	2252; 364;	1224; 1449;	1283; 710;	1855; 907;	1542; 2903;		1,86	1,387	1,572	303	673	741
33,03	1698; 1168;	2375; 443;	1333; 1666;	1283; 731;	1842; 1183;	1512; 2900;		1,869	1,19	1,632	144	911	702

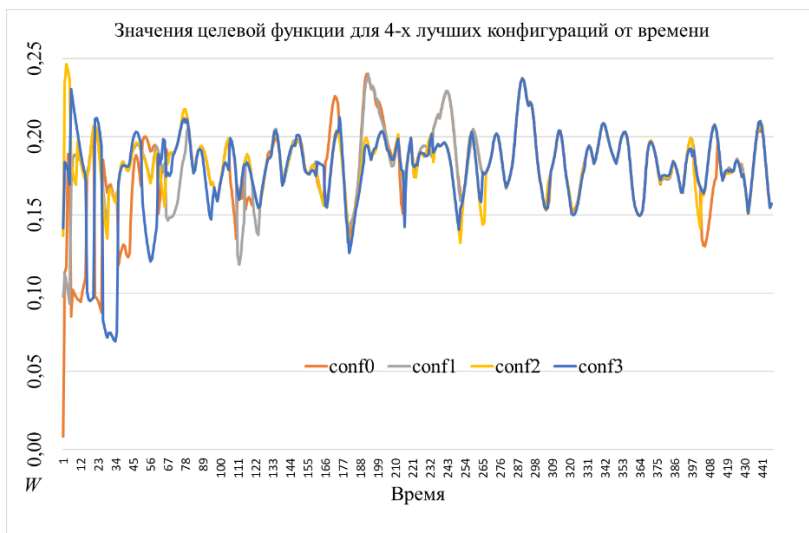
Моделирование и добавление новых данных в таблицы продолжается до тех пор, пока пользователь не прервет программу. На рис. 3 можно заметить, что предложенное расположение точек доступа обеспечивает покрытие значительной части маршрутов, а различия между конфигурациями становятся значительно меньше.



*Рис. 3. Результат моделирования*

Изменения уровня доступа в зависимости от времени можно отследить на рис. 4. Видно, что первые две минуты моделирования различия между конфигурациями заметны, а значение целевой функции быстро растет. Затем различия между конфигурациями становятся минимальными, и изменения в уровне доступа больше связаны с движением абонентов, чем с изменением положения сетевых устройств.

Поиск оптимального значения целевой функции ведется непрерывно, моделирование не сводится к конкретному результату, а, скорее, служит инструментом для иллюстрации возможных мест расположения сетевых устройств и обеспечиваемого уровня доступа.



*Рис. 4. Изменения значений целевой функции  
в зависимости от времени*

## **5. Выводы и оценка результатов**

В связи с непрерывным движением абонентов решение не сводится к конкретному минимуму, точки доступа будут продолжать свое движение в зависимости от расположения абонентов. Но хотя конечное решение не будет предложено, полученные результаты имеют практическую ценность. Перемещение сетевых устройств спустя первые несколько минут моделирования ограничивается некоторой областью, решение о размещении устройств в конкретных координатах может быть принято эмпирически с учетом полученных результатов и технических возможностей, не отраженных в модели. Возможно проведение дополнительного моделирования с фиксированными координатами сетевых устройств для расчета обеспечиваемого уровня доступа и выбора из нескольких вариантов. Программа может быть использована для решения о целесообразности привлечения дополнительного оборудования путем проведения моделирования с другим количеством сетевых устройств. Если сетевое оборудование возможно передвигать, может быть принято ре-

шение о траектории его перемещения. Например, это может быть автомобиль с готовым к работе сетевым оборудованием, и его маршрут будет организован с учетом результатов моделирования и технических возможностей, например, наличия дорог. Достоинствами модели являются также низкая вычислительная сложность и, как следствие, быстрое получение результатов.

В ходе дальнейшей работы предполагается реализовать возможность загрузки рельефа с помощью карт и учитывать помехи при распространении сигнала. Используемые функции в некоторых экспериментах могут привести к наличию абонентов, не обеспеченных связью, находящихся на удаленном маршруте, так как в соответствии с выбранной функцией выгоднее увеличить надежность связи для близкорасположенных абонентов. Будут выполнены дополнительные исследования и эксперименты по эффективности использования других целевых функций.

Пользовательский интерфейс ограничен просмотром конфигураций в 2d- или 3d-представлении и отображением таблиц результатов. Все настройки и ввод исходных данных осуществляются в конфигурационном файле, что требует внимания и некоторого опыта работы. Предполагается доработка интерфейса.

Программное приложение было разработано в среде Microsoft Visual Studio на языке C++. Выбор среды разработки обоснован ее универсальностью и широкими возможностями, а также популярностью, надежностью, удобным интерфейсом и большим количеством средств отладки кода.

Одно из направлений практического применения сетей описанного вида – поисково-спасательные, исследовательские или спортивные мероприятия, проводимые в горной местности. В связи с особенностями рельефа плотность населения в таких областях очень невысока, что делает нецелесообразной установку вышек сотовой связи, достаточных для покрытия территории сетью мобильной связи и доступом к сети Интернет. Разработанный программный продукт ориентирован на быструю оценку возможности использования сети в конкретных условиях, что позволит чаще и эффективнее применять беспроводные сети быстрого развертывания. Это приведет к уменьшению числа мероприятий, не обеспеченных связью, улучшению связи и ко-

ординации как между участниками, так и между ними и спасательными службами, а в случае необходимости – экспертами. Все вышеперечисленное позволит сократить время ликвидации последствий чрезвычайных ситуаций в малонаселенных районах и повысить безопасность проводимых мероприятий.

## **6. Заключение**

В статье рассмотрен программный продукт, разработанный авторами для быстрого решения проблемы проектирования беспроводных сетей связи. Задача формулируется как поиск мест размещения сетевых устройств с известными характеристиками для обеспечения связью абонентов, перемещающихся по заданным маршрутам, при невозможности полного покрытия территории сетью связи. Для решения задачи используются известные методы моделирования (генетические алгоритмы, имитационное моделирование), их применение к данной задаче и все полученные результаты являются новыми.

## **Литература**

1. БАЛИЦКИЙ В.С. и др. *Мобильная станция видеомониторинга и связи*. – 2010.
2. БЛИНОВА О.В., ВАСЬКОВСКИЙ С.В., ФАРХАДОВ М.П. *Оценка надежности системы связи с подвижными узлами // Датчики и системы*. – 2018. – №5. – С. 3–8.
3. БЛИНОВА О.В., ВАСЬКОВСКИЙ С.В., ФАРХАДОВ М.П. *Взаимосвязь подвижных абонентов и стационарных узлов связи при известных характеристиках движения // Датчики и системы*. – 2017. – №3. – С. 3–8.
4. БЛИНОВА О.В., ВАСЬКОВСКИЙ С.В., ВОРОНЦОВ Ю.А., ФАРХАДОВ М.П. *Информационная система с подвижными узлами связи // Датчики и системы*. – 2015. – №12(198). – С. 24–28.
5. БОРИСОВ Ю.Ю. *Система с ретрансляцией емкости как средство обеспечения широкополосного доступа в малых населенных пунктах, на автомобильных и железных доро-*

- гах // Итоги Второй конференции “Технологии мобильной и беспроводной связи. Тренды и перспективы”. – 2014.
6. БУТРОВ В.А. и др. *Некоторые аспекты построения и организации сетей спутниковой связи специального назначения* // Вопросы электромеханики. Труды ВНИИЭМ. – 2008. – Т. 107. – С. 55–59.
  7. ВИШНЕВСКИЙ В.М., КРИШНАМУРТИ А., КОЗЫРЕВ Д.В., ЛАРИОНОВ А.А., ИВАНОВ Р.Е. *Методы исследования и проектирования широкополосных беспроводных сетей вдоль протяженных транспортных магистралей* // T-Comm. – 2015. – №5. – С. 9–15.
  8. ГЛУЩЕНКО Е.А., ЮРИН И.В. *Применение современных интернет-технологий на водном транспорте* // Вестник государственного университета морского и речного флота им. адмирала СО Макарова. – 2011. – №1(9). – С. 54–58.
  9. ИВАНОВ Р.Е., МУХТАРОВ А.А., ПЕРШИН О.Ю. *Задача оптимального размещения заданного множества базовых станций беспроводной сети связи с линейной топологией* // Автоматизация, телемеханизация и связь в нефтяной промышленности. – 2019. – №4. – С. 39–45.
  10. КУН Г.Р., ЧИКИН А.Е. *Перспективы развития приборостроения в условиях импортозамещения на базе АО "НПФ "Микран"* // Метрология в радиоэлектронике. – 2016. – С. 5–8.
  11. AGRAWAL S., MISRA A.K., TYAGI N. *Seamless VANET connectivity through heterogeneous wireless network on rural highways* // Proc. of the Second Int. Conf. on Information and Communication Technology for Competitive Strategies. – 2016. – P. 1–5.
  12. ANDRUSENKO J., KASCH W.T., WARD J.R. *Wireless network modeling and simulation tools for designers and developers* // IEEE Communications Magazine. – 2009. – Vol. 47, No. 3. – P. 120–127.
  13. BUDIMIR D., SHELKOVNIKOV B.N. *CAD for broadband wireless access design* // Proc. of the IEEE 5th Int. Conf. on Telecommunications in Modern Satellite, Cable and Broadcasting Service (TELSIKS-2001). (Cat. No. 01EX517). – IEEE, 2001. – Vol. 2. – P. 525–528.



14. DALMASSO I. et al. *WiMAX networks for emergency management based on UAVs* // IEEE First AESS European Conference on Satellite Telecommunications (ESTEL-2012). – IEEE, 2012. – P. 1–6.
15. DOWN R. SCHEINERT S. *Internet base station: US patent 7117015*. – 2006.
16. HANNIKAINEN M., HAMALAINEN T.D., VANHATUPA T. *Genetic algorithm to optimize node placement and configuration for WLAN planning* // IEEE 4th Int. Symposium on Wireless Communication Systems. – IEEE, 2007. – P. 612–616.
17. KALIKA L. et al. *Node placement method within a wireless network, such as a wireless local area network* : US patent 7539495. – 2009.
18. KASPRZYK M.Z., OTTO K. *Wireless Communication Apparatus for Emergency Situations* : US patent 13708911. – 2013.
19. <https://landcomm.ru/dokumentacija/11383/>.

## **DESIGN RAPID DEPLOYMENT COMMUNICATION NETWORKS USING THE NETWORK CONFIGURATION ANALYSIS SOFTWARE APPLICATION**

**Olga Blinova**, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, research associate (blinova\_olga\_v@mail.ru).

**Sergey Vaskovsky**, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand.Sc., senior researcher (v63v@yandex.ru).

**Yaroslav Rykov**, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, control engineer (yarykov@gmail.com).

*Abstract: A new approach to the design of rapid deployment wireless communication networks is proposed by the conditions when it is impossible to cover fully the territory with a network. The essence of this approach is to determine the optimal locations of specified sets of network equipment prepared in advance. When performing the search, the specified routes of subscribers and their speed are taken into account. The target function uses the time- depended total network access level for all subscribers. The main existing methods of designing wireless networks, their application possibilities and disadvantages when used in a certain class of tasks (rapid deployment, incomplete coverage) are also considered. For practical application of the considered approach, a software product has been created that implements the developed search procedure using simulation and genetic algorithms. The algorithm of the software product operation is considered and examples of simula-*

*tion results are given. The advantages and disadvantages of the proposed solution are analyzed. The problem of placing network devices in the specified conditions is solved for the first time; the software product created has no analogues. The results obtained are applicable for rapid design of wireless communication networks in areas where there is no mobile network coverage, for example, to provide communication for rescue operations or tourist routes in mountainous areas.*

Keywords: wireless communication networks, communication networks with incomplete coverage, providing communication for rescue operations, rapid deployment communication networks.

УДК 004.7 + 004.9

ББК 32.972.5

DOI: 10.25728/ubs.2021.90.6

*Статья представлена к публикации  
членом редакционной коллегии Я.И. Квинто.*

*Поступила в редакцию 15.10.2020.*

*Опубликована 31.03.2021.*