

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«СИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ТЕЛЕКОММУНИКАЦИЙ И ИНФОРМАТИКИ»

Вестник СибГУТИ № 2 (46) 2019

Выпускается ежеквартально, выходит с 2007 г.

Председатель редакционного совета В. Г. Беленький, к.т.н., доц.

Заместитель председателя редакционного совета А. Н. Фионов, д.т.н., проф.

Редакционный совет:

В. П. Бакалов, д.т.н., проф.	В. И. Носов, д.т.н., проф.
В. Б. Барахнин, д.т.н., доц.	В. П. Петров, д.т.н., проф.
В. Ю. Васильев, д.х.н.	С. В. Поршнев, д.т.н., проф.
В. Н. Васюков, д.т.н., проф.	А. С. Родионов, д.т.н., с.н.с.
Н. И. Горлов, д.т.н., проф.	А. И. Романенко, д.ф.-м.н., проф.
Н. Л. Казначеева, д.э.н., доц.	Б. Я. Рябко, д.т.н., проф.
В. С. Канев, д.т.н., доц.	И. И. Рябцев, д.ф.-м.н.
А. И. Карпович, д.э.н., проф.	Э. Сименс, д.т.н.
Б. И. Крук, к.т.н., проф.	О. В. Стукач, д.т.н., проф.
В. В. Лебедянцев, д.т.н., проф.	В. К. Трофимов, д.т.н., проф.
А. В. Лихачёв, д.т.н.	А. И. Фалько, д.т.н., проф.
С. Н. Мамоиленко, д.т.н., доц.	С. В. Федоренко, д.т.н., доц.
В. Г. Мартынец, д.ф.-м.н.	А. М. Федотов, д.ф.-м.н., чл.-корр. РАН
А. Б. Мархасин, д.т.н., проф.	А. Г. Черевко, к.ф.-м.н., доц.
О. Г. Мелентьев, д.т.н., проф.	С. В. Шидловский, д.т.н.
Р. В. Мещеряков, д.т.н., проф.	Ю. И. Шокин, д.ф.-м.н., акад. РАН
И. Г. Неизвестный, д.ф.-м.н., чл.-корр. РАН	В. П. Шувалов, д.т.н., проф.

Редакция:

А. Н. Фионов (главный редактор), М. Ю. Галкина (заведующая редакцией),
Н. А. Двуреченская (технический и литературный редактор, компьютерная вёрстка),
Т. А. Алфёрова (лингвист-корректор), Е. Л. Грозовская (секретарь)

Адрес редакции

630102, г. Новосибирск, ул. Кирова, д. 86

e-mail: vestnik@sibsutis.ru

Информация о журнале доступна в сети Internet по адресу <http://vestnik.sibsutis.ru>.

Журнал включён в Перечень ВАК российских рецензируемых научных журналов, в которых должны быть опубликованы основные научные результаты диссертаций на соискание учёных степеней доктора и кандидата наук.

Журнал зарегистрирован Федеральной службой по надзору за соблюдением законодательства в сфере массовых коммуникаций и охране культурного наследия, свидетельство о регистрации ПИ № ФС77-25835 от 29.09.2006. ISSN 1998-6920. Подписной индекс в объединённом каталоге «Пресса России» – 82519.

Отпечатано в издательском центре СибГУТИ. Бумага офсетная, формат А4. Тираж 300 экз.

© СибГУТИ, 2019 г.

СОДЕРЖАНИЕ

В. П. Бакалов, Е. А. Субботин

Масштабирование нерекурсивных цифровых фильтров с нерекурсивными
латеральными связями 3

Л. С. Петрова, Е. В. Заец

Математическое моделирование переходных термических процессов с учётом
явлений тепловой релаксации и термического демпфирования 13

В. Г. Арсентьев, Г. И. Криволапов

Измерение разности фаз гармонических сигналов в линейных трактах
с квазибелым шумом 21

А. И. Ицксон

Ретроанализ менеджмента эксплуатации систем электросвязи 32

М. Г. Курносов

Анализ и оптимизация конвейерных алгоритмов широкополосной передачи
стандарта МРІ 43

В. Р. Губкина

Теоретико-множественная модель для решения задачи системного анализа и
синтеза в телекоммуникациях 57

Е. М. Ильин, Д. А. Репников, В. Ю. Савостьянов, О. Ф. Самарин,

А. И. Полубехин, А. Г. Черевко

Режимы функционирования многофункциональной бортовой РЛС БЛА
малой и средней дальности 68

С. И. Безденежных, С. Г. Брайткрайц, Д. А. Репников, В. И. Флегонтов,

А. Г. Черевко

Эволюционно-технологический подход в управлении требованиями
к робототехническим комплексам военного назначения 87

В. П. Петров, И. Ю. Якушев

Современные технологии в системе МІМО 94

Масштабирование нерекурсивных цифровых фильтров с нерекурсивными латеральными связями

В. П. Бакалов, Е. А. Субботин

В статье излагаются способы масштабирования цифровых фильтров (ЦФ) с латеральными обратными связями (ЛОС). Рассматриваются модели четырех типов ЦФ с ЛОС при их нерекурсивной реализации. Масштабирование фильтров осуществляется с использованием импульсной характеристики h_i и АЧХ цепи. Полученные результаты позволяют, с одной стороны, предотвратить переполнение разрядной сетки ЦФ с ЛОС, а с другой – существенно повысить его помехоустойчивость и эффективность.

Ключевые слова: нерекурсивные цифровые фильтры, нерекурсивные латеральные обратные связи, масштабирование, ограничение максимума сигнала, ограничение энергии сигнала, ограничение максимума АЧХ цепи.

1. Введение

Как известно, при цифровой обработке сигналов сложение чисел с фиксированной запятой может приводить в цифровом фильтре (ЦФ) к такому опасному явлению, как переполнение [1, 2]. Чтобы в ЦФ не возникало переполнения, применяют в определенных точках фильтра процедуру масштабирования множителей, которая позволяет, с одной стороны, предотвратить переполнение, а с другой – сохранить максимально возможное отношение сигнала к уровню шума округления в ЦФ.

Особенно важна проблема масштабирования в ЦФ с латеральными обратными связями (ЛОС), использование которых позволяет существенно повысить помехоустойчивость и эффективность цифровых фильтров [4–6]. В статье [8] рассматривались вопросы масштабирования ЦФ с ЛОС при их рекурсивной реализации.

В данной работе рассматриваются вопросы расчета масштабных множителей для четырех типов ЦФ с ЛОС при их нерекурсивной реализации по условиям ограничения максимума сигнала, максимума энергии и максимума АЧХ цепи.

2. Цифровой фильтр с ЛОС 1-го типа

Рассмотрим модель двухканального одноуровневого ЦФ с ЛОС 1-го типа [3], представленную на рис. 1.

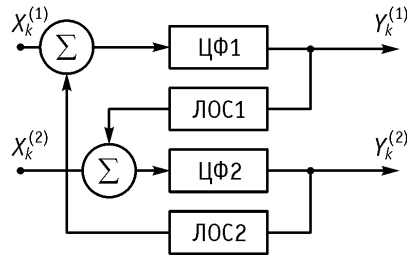


Рис. 1. Двухканальная система с ЛОС. Тип № 1

Будем считать, что ЦФ и ЛОС реализуются в виде нерекурсивных цепей n -го порядка.

Положим, что шумы квантования и округления на выходе каждого из умножителей и на выходе аналого-цифровых преобразователей (АЦП) некоррелированы между собой и некоррелированы с входными последовательностями каждого канала. Причем источники шума $e_1 \dots e_n$ имеют одинаковую дисперсию и определяются уравнениями

$$\sigma_k^2 = \sigma^2(kT) = \Delta^2/12, \quad (1)$$

а дисперсия шума на выходе ЦФ – уравнением:

$$\sigma^2 = \frac{\Delta^2}{12} \sum_{k=1}^n h_k^2, \quad (2)$$

где Δ определяется числом используемых разрядов; h_k – импульсная характеристика ЦФ.

С учетом вышеизложенного на рис. 2 изображена шумовая модель ЦФ с ЛОС этого типа, где e_0 – шумовой сигнал, учитывающий шум квантования входного сигнала; e_1, e_n – шумовой сигнал, учитывающий квантование сигналов в умножителях.

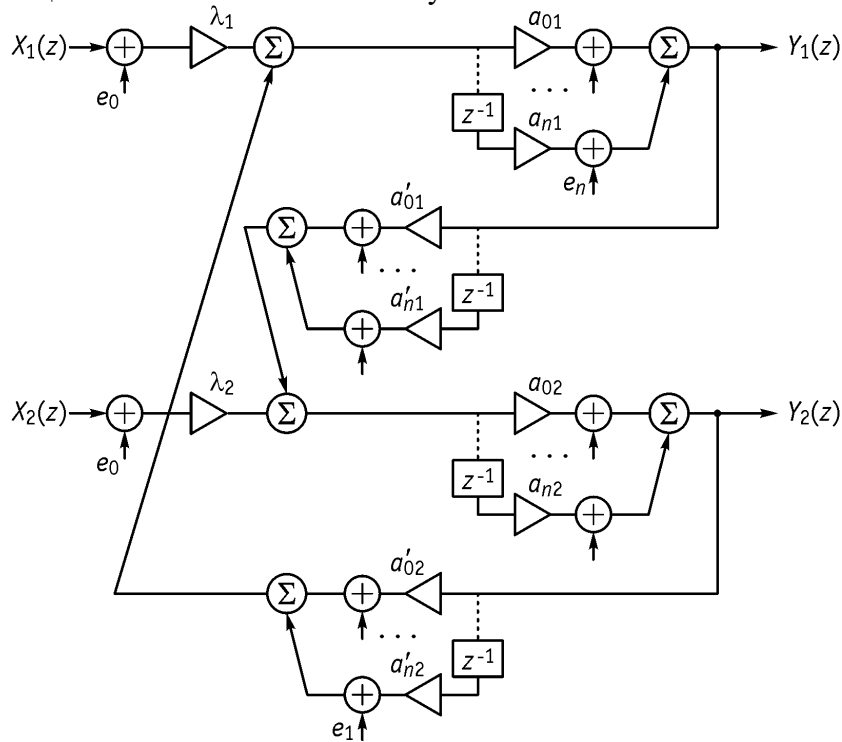


Рис. 2. Модель НЦФ с НЛОС 1-го типа

Передаточная функция нерекурсивного цифрового фильтра (НЦФ1) определяется уравнением

$$H_1(z) = \sum_{n=0}^{N_1} a_{n1} z^{-n}, \quad (3)$$

где N_1 – количество умножителей прямой связи НЦФ1.

Аналогично определяется передаточная функция НЦФ2:

$$H_2(z) = \sum_{n=0}^{N_2} a_{n2} z^{-n}, \quad (4)$$

где N_2 – количество умножителей прямой связи НЦФ2.

Передаточная функция нерекурсивного фильтра ЛОС1 определяется уравнением:

$$H_{\text{ЛОС1}}(z) = \sum_{n=0}^{N'_1} a'_{n1} z^{-n}, \quad (5)$$

где N'_1 – количество умножителей прямой связи нерекурсивного фильтра ЛОС1.

Передаточная функция нерекурсивного фильтра ЛОС2 определяется уравнением:

$$H_{\text{ЛОС2}}(z) = \sum_{n=0}^{N'_2} a'_{n2} z^{-n}, \quad (6)$$

где N'_2 – количество умножителей прямой связи нерекурсивного фильтра ЛОС2; a_{n1}, a_{n2} – коэффициенты, характеризующие веса нерекурсивной части ЦФ1 и ЦФ2 соответственно; a'_{n1}, a'_{n2} – коэффициенты, характеризующие веса нерекурсивной части ЛОС1 и ЛОС2 соответственно.

Определим масштабные коэффициенты данного фильтра с использованием максимума сигнала, максимума энергии сигнала и максимума АЧХ цепи.

Положим, что импульсная характеристика ЦФ с ЛОС (или его части) равна $h_i(n)$, тогда выходной сигнал фильтра (или его части) $y_i(n)$ будет ограничен величиной

$$\max |y_i(n)| \leq \max |x(n)| \sum_{n=0}^{\infty} |h_i(n)|. \quad (7)$$

Если принять, что $\max |x(n)| \leq 1$, то необходимым условием отсутствия переполнения является:

$$\sum_{n=0}^{\infty} |h_i(n)| \leq 1. \quad (8)$$

Для обеспечения данного условия введем масштабные коэффициенты λ_i :

$$\lambda_i \sum_{n=0}^{\infty} |h_i(n)| \leq 1. \quad (9)$$

Откуда находим λ_i по условию ограничения максимума сигнала:

$$\lambda \leq \frac{1}{\sum_{n=0}^{\infty} |h_i(n)|}, \quad (10)$$

где $h_i(n)$ – импульсная характеристика участка ЦФ от входа до выхода i -го сумматора.

Для определения масштабного множителя по условию (10) найдем импульсные характеристики 1-го канала для данного типа ЦФ с ЛОС.

Для первого сумматора передаточная функция определяется уравнением:

$$H''_2(z) = H_2(z) H_{\text{ЛОС2}}(z), \quad (11)$$

где $H_2(z)$ и $H_{\text{ЛОС2}}(z)$ определяются формулами (4) и (6).

Для второго сумматора импульсная характеристика соответствует передаточной функции всего фильтра $H_1(z)$, определяемой (3).

После нахождения передаточных функций $H_2''(z)$ и $H_1(z)$ импульсные характеристики определяются с помощью обратного z -преобразования [7]. В частности, для первого сумматора импульсная характеристика будет равна:

$$h_{1\Sigma}(n) = \frac{1}{2\pi j} \oint_{|z|=1} H_2''(z) z^{n-1} dz, \quad (12)$$

для второго сумматора:

$$h_{2\Sigma}(n) = \frac{1}{2\pi j} \oint_{|z|=1} H_1(z) z^{n-1} dz, \quad (13)$$

где $H_1(z)$, $H_2''(z)$ определяются формулами (3), (11).

После нахождения импульсных характеристик $h_i(n)$ масштабные коэффициенты для первого $\lambda_{1\Sigma}$ и второго $\lambda_{2\Sigma}$ сумматоров определяются согласно (10) и из них выбирается наименьший: $\lambda = \min \lambda_{\Sigma}$.

Аналогичным образом можно использовать импульсные характеристики h_i при масштабировании ЦФ с ЛОС при ограничении энергии сигнала [8]:

$$\lambda \leq \frac{1}{\sqrt{\sum_{n=0}^{\infty} h_i^2(n)}}, \quad (14)$$

где $h_i(n)$ определяется согласно (12), (13).

При расчете масштабных множителей по условию ограничения усиления цепи необходимо задаться неравенством

$$\lambda = \frac{1}{\max H(\omega)}, \quad (15)$$

где $H(\omega) = |He^{j\omega T}|$.

При этом максимальное значение АЧХ цепи $H_{\max}(\omega)$ определяется на частоте, равной собственной частоте нуля или полюса передаточной функции $H(z)$ [7].

Учитывая, что масштабирование обычно приводит к существенному уменьшению отношения сигнал/шум на выходе ЦФ, определим дисперсию шума на выходе ЦФ с ЛОС для различных схем реализации.

На основании шумовой модели (рис. 2) и уравнений (1), (2) определим дисперсию шума на выходе НЦФ с НЛОС с учетом масштабирования:

$$\begin{aligned} \sigma_{\text{вых}}^2 = & \frac{\lambda_1^2 (2^{-b_{\text{АЦП}}})^2}{12} \left[\sum_{n=0}^{\infty} h_1^2(n) + \sum_{n=0}^{\infty} h_3^2(n) \right] + \\ & + \frac{(2^{-b_{\text{умн}}})^2}{12} \left[N_2' \sum_{n=0}^{\infty} h_1^2(n) + N_1' \sum_{n=0}^{\infty} h_4^2(n) + N_2 \sum_{n=0}^{\infty} h_6^2(n) + N_1 \right], \end{aligned} \quad (16)$$

где $h_1(n) \div H_1(z)$; $h_3(n) \div H_1(z) \cdot H_{\text{ЛОС1}}(z) \cdot H_2(z)$;

$h_4(n) \div H_2(z) \cdot H_{\text{ЛОС2}}(z) \cdot H_1(z)$; $h_6(n) \div H_{\text{ЛОС2}}(z) \cdot H_1(z)$.

Из формулы (16) следует, что дисперсия шума на выходе нерекурсивного ЦФ с нерекурсивной ЛОС будет определяться суммой дисперсий отдельных составляющих, причем наибольший вклад в дисперсию шума на выходе вносят множители, хотя вклад этот существенно меньше, чем в рекурсивных ЦФ с рекурсивной ЛОС [8].

3. Цифровой фильтр с ЛОС 2-го типа

Рассмотрим масштабирование ЦФ с ЛОС 2-го типа (рис. 3).

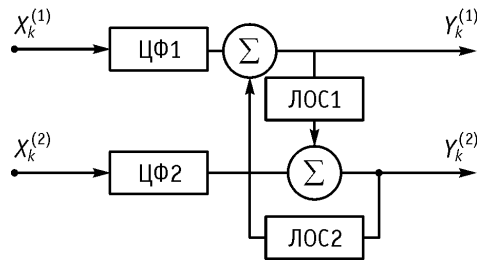


Рис. 3. Двухканальная система с ЛОС. Тип № 2

На рис. 4 представлена шумовая модель этого фильтра при нерекурсивной реализации.

Для этого фильтра передаточная функция первого сумматора равна нулю, а второго сумматора – определяется передаточной функцией ЦФ1 и передаточной функцией $H_2''(z)$ согласно (11).

При этом импульсная характеристика второго сумматора будет равна:

$$h_{2\Sigma}(n) = h_1(n) + h_2''(n) \div (H_1(z) + H_2''(z)), \quad (17)$$

где $H_1(z)$ определяется (3), а $H_2''(z)$ – формулой (11).

После нахождения $h_{2\Sigma}(n)$ по формуле (10) находим масштабные коэффициенты $\lambda_{1\Sigma}$ и $\lambda_{2\Sigma}$ по условию максимума сигнала и выбираем из них наименьший.

Аналогичным образом определяются масштабные множители по условиям (14) и (15).

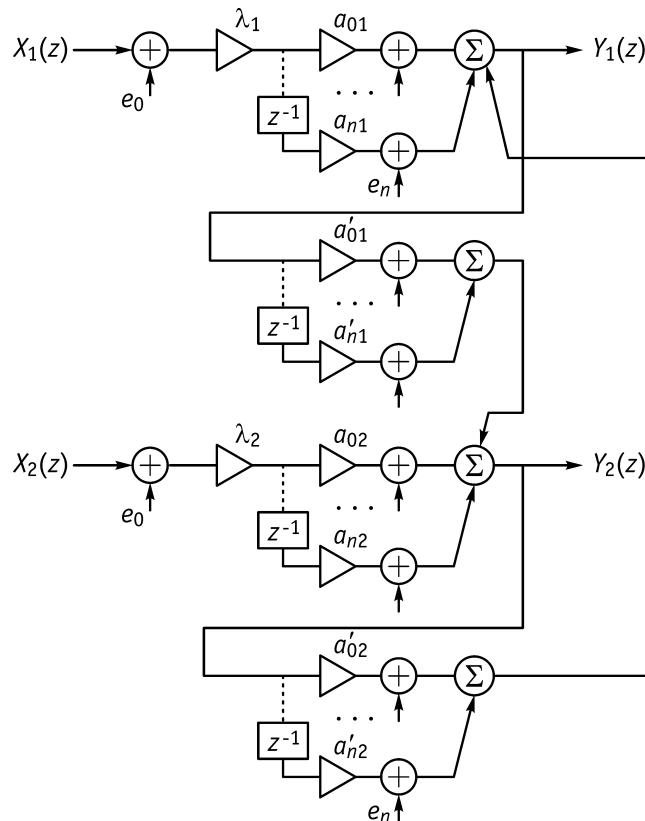


Рис. 4. Модель НЦФ с НЛОС 2-го типа

Дисперсия шума на выходе НЦФ с НЛОС 2-го типа после масштабирования определяется уравнением:

$$\sigma_{\text{ВЫХ}}^2 = \frac{\lambda_1^2 (2^{-b_{\text{АЦП1}}})^2}{12} \left[\sum_{n=0}^{\infty} h_1^2(n) + \sum_{n=0}^{\infty} h_5^2(n) \right] + \frac{\lambda_2^2 (2^{-b_{\text{АЦП2}}})^2}{12} \sum_{n=0}^{\infty} h_4^2(n) + \frac{(2^{-b_{\text{УМН}}})^2}{12} \left[(N_1) \sum_{n=0}^{\infty} h_7^2(n) + (N_2 + N'_1) \sum_{n=0}^{\infty} h_9^2(n) + (N_1 + N'_2) \right], \quad (18)$$

где $h_1(n) \div H_1(z)$; $h_5(n) \div H_{\text{ЛОС1}}(z) \cdot H_{\text{ЛОС2}}(z) \cdot H_1(z)$;

$h_4(n) \div H_2(z) \cdot H_{\text{ЛОС2}}(z)$; $h_7(n) \div H_{\text{ЛОС1}}(z) \cdot H_{\text{ЛОС2}}(z)$; $h_9(n) \div H_{\text{ЛОС2}}(z)$.

Анализ уравнения (18) показывает, что, несмотря на увеличение дисперсии от шумов АЦП, дисперсия шума на выходе НЦФ не так сильно увеличивается по сравнению с 1-й схемой за счет дополнительного увеличения путей прохождения шума и суммарной дисперсии шума умножителей и существенно меньше дисперсии шума в рекурсивном варианте ЦФ с ЛОС.

4. Цифровой фильтр с ЛОС 3-го типа

Рассмотрим ЦФ с ЛОС 3-го типа (рис. 5), шумовая модель которого изображена на рис. 6.

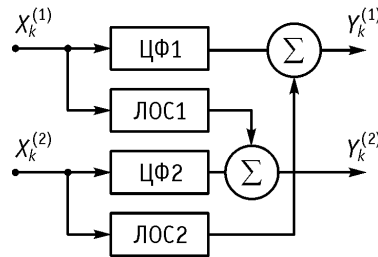


Рис. 5. Двухканальная система с ЛОС. Тип № 3

Передаточная характеристика первого сумматора равна нулю, а второго – определяется передаточной функцией $H_1(z)$ (3) и передаточной функцией ЛОС2 согласно (6).

После нахождения $h_{2\Sigma}(n) \div (H_1(z) + H_{\text{ЛОС2}}(z))$ множители λ определяются по формулам (10), (14), (15) в зависимости от условий масштабирования.

Дисперсия шума на выходе НЦФ с НЛОС 3-го типа после масштабирования определяется уравнением:

$$\sigma_{\text{ВЫХ}}^2 = \frac{\lambda_1^2 (2^{-b_{\text{АЦП1}}})^2}{12} \sum_{n=0}^{\infty} h_1^2(n) + \frac{\lambda_2^2 (2^{-b_{\text{АЦП2}}})^2}{12} \sum_{n=0}^{\infty} h_4^2(n) + \frac{(2^{-b_{\text{УМН}}})^2}{12} [(N_1 + N'_2) \cdot 1], \quad (19)$$

где $h_1(n) \div H_1(z)$; $h_4(n) \div H_{\text{ЛОС2}}(z)$.

Анализ уравнения (19) показывает, что шумы АЦП в схемах 1 и 3 при одинаковых разрядных сетках примерно одинаковы и меньше, чем во 2-й схеме, а шумы умножителей в 3-й схеме существенно меньше, чем в предыдущих схемах, и тем более, чем в схемах рекурсивных ЦФ, за счет уменьшения путей прохождения шума, что приводит к уменьшению суммарной дисперсии шума на выходе НЦФ с НЛОС 3-го типа.

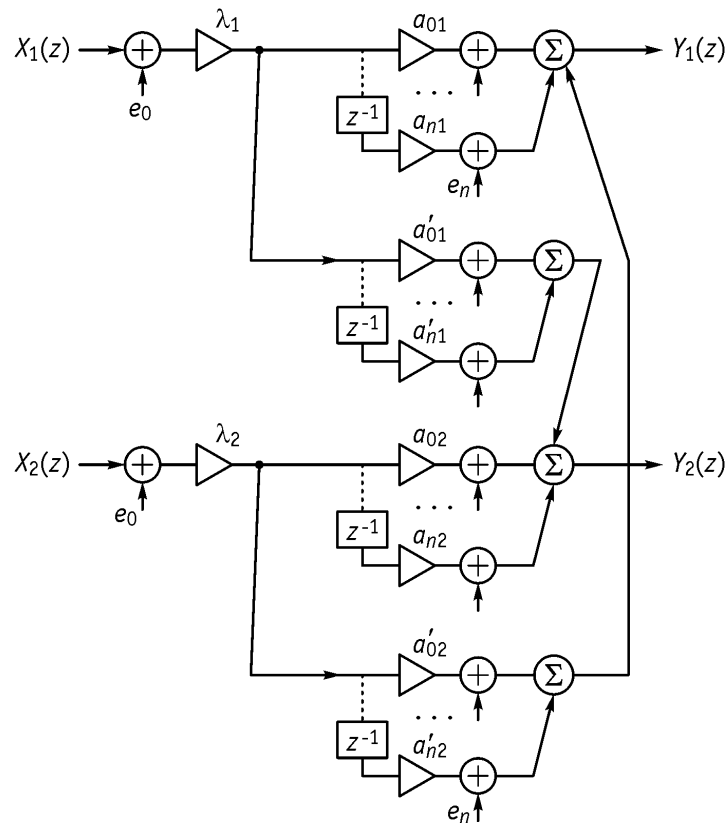


Рис. 6. Модель НЦФ с НЛОС 3-го типа

5. Цифровой фильтр с ЛОС 4-го типа

В заключение рассмотрим схему ЦФ с ЛОС 4-го типа (рис. 7), шумовая модель которого изображена на рис. 8.

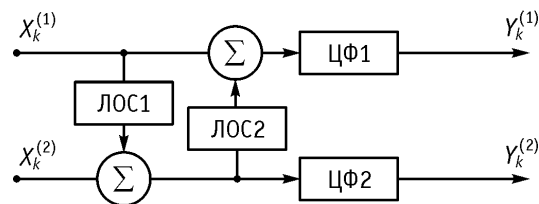


Рис. 7. Двухканальная система с ЛОС. Тип № 4

Импульсная характеристика первого сумматора определяется передаточной функцией ЛОС $H_{\text{ЛОС2}}(z)$ (6).

Импульсная характеристика второго сумматора определяется передаточной функцией $H_1(z)$ (3).

После нахождения импульсных характеристик для 1-го и 2-го сумматоров определяются соответствующие масштабные множители согласно (10), (14), (15). Аналогичным образом находятся масштабные множители для 2-го канала.

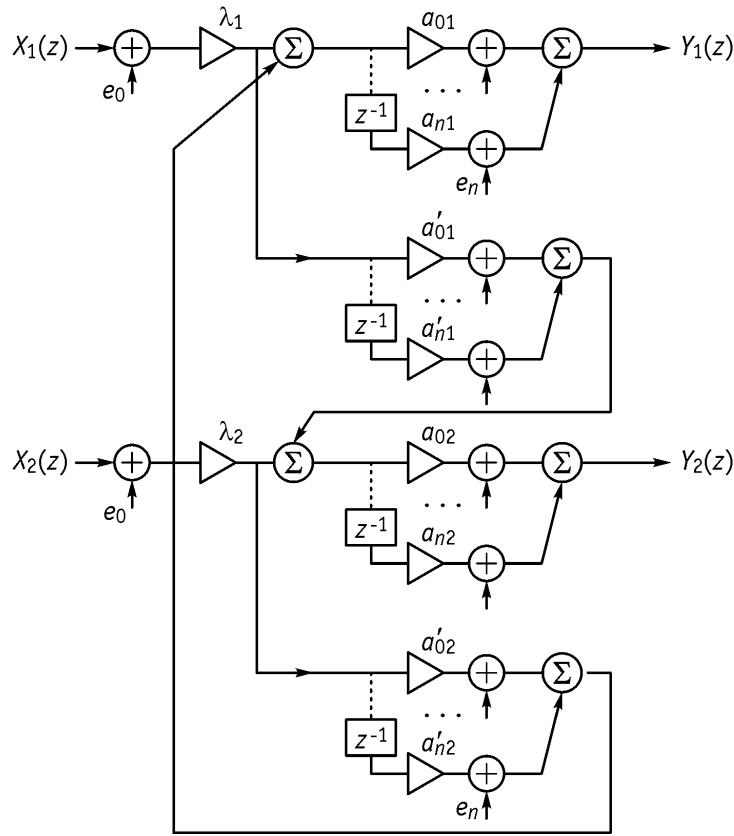


Рис. 8. Модель НЦФ с НЛОС 4-го типа

Дисперсия шума на выходе ЦФ с ЛОС 4-го типа после масштабирования определяется уравнением:

$$\sigma_{\text{вых}}^2 = \frac{\lambda_1^2 (2^{-b_{\text{АЦП1}}})^2}{12} \left[\sum_{n=0}^{\infty} h_1^2(n) + \sum_{n=0}^{\infty} h_5^2(n) \right] + \frac{\lambda_2^2 (2^{-b_{\text{АЦП2}}})^2}{12} \sum_{n=0}^{\infty} h_4^2(n) + \frac{(2^{-b_{\text{УМН}}})^2}{12} \left[N_2' \sum_{n=0}^{\infty} h_1^2(n) + N_1' \sum_{n=0}^{\infty} h_4^2(n) + N_2' \sum_{n=0}^{\infty} h_5^2(n) + N_1 \right], \quad (20)$$

где $h_1(n) \div H_1(z)$; $h_4(n) \div H_1(z) \cdot H_{\text{ЛОС2}}(z)$; $h_5(n) \div H_{\text{ЛОС1}}(z) \cdot H_{\text{ЛОС2}}(z) \cdot H_1(z)$.

Анализ уравнения (20) показывает, что в данной схеме дисперсия шума АЦП, как и в рекурсивном ЦФ с рекурсивной ЛОС, при прочих равных условиях больше, чем дисперсия шума АЦП в схеме № 1 и схеме № 3, и одинакова с дисперсией шума АЦП в схеме № 2. Дисперсия шума от умножителей в схеме № 4 значительно выше, чем в схемах первых трех моделей, что приводит к существенному увеличению дисперсии на выходе НЦФ с НЛОС этого типа в целом.

Анализируя рассмотренные шумовые модели следует отметить, что в схеме 4-го типа шумы от источников a'_{n1} и a'_{n2} имеют, как и при рекурсивной реализации, циклические пути $H_{\text{ЛОС1}}(z) \cdot H_{\text{ЛОС2}}(z) \cdot H_2(z)$ и $H_{\text{ЛОС1}}(z) \cdot H_{\text{ЛОС2}}(z) \cdot H_1(z)$, то есть дисперсия шума в такой схеме будет увеличиваться. Из рассмотренных схем также следует, что, как и для рекурсивных ЦФ с рекурсивной ЛОС, наиболее оптимальной с точки зрения минимума дисперсии шума на выходе является 3-я схема. Причем для всех рассмотренных схем дисперсия шума на выходе НЦФ с НЛОС будет существенно меньше дисперсии шума на выходе соответствующих схем рекурсивных ЦФ с рекурсивной ЛОС. Это объясняется существенно меньшим количеством умножителей в НЦФ с НЛОС, вносящих наибольший вклад в результирующий шум на выходе фильтра.

Если сравнивать по помехоустойчивости масштабируемые нерекурсивные ЦФ с нерекурсивными ЛОС, то лучший результат дает схема 3, затем схемы 2 и 4, а наименьший выигрыш дает схема 1.

Полученные результаты по масштабированию ЦФ с ЛОС можно обобщить и на n -канальную систему с ЛОС. Причем в зависимости от вида помех могут изменяться характер ЛОС, структура охвата каналов латеральными связями (каждый канал с каждым, по группам мешающих параметров и т.д.), подбор глубины обратной связи.

В заключение следует отметить, что процесс масштабирования в цифровых фильтрах, как отмечалось, приводит к существенному уменьшению отношения сигнал/шум на выходе фильтра по сравнению с отношением сигнал/шум на входе, что, в свою очередь, ухудшает помехоустойчивость ЦФ в целом [1]. Использование же принципа ЛОС, как показано в [3], позволяет существенно повысить помехоустойчивость ЦФ даже при отношении сигнал/помеха на входе меньше единицы. Таким образом, при масштабировании ЦФ с ЛОС не происходит уменьшения отношения сигнал/шум на выходе фильтра, что существенно повышает его помехоустойчивость по сравнению с традиционными методами масштабирования ЦФ.

6. Заключение

Исследованы вопросы масштабирования нерекурсивных двухканальных одноуровневых цифровых фильтров (ЦФ) с нерекурсивными латеральными обратными связями (ЛОС) по условиям максимума сигнала, максимума энергии сигнала и максимума АЧХ.

Определены масштабные коэффициенты по условиям ограничения максимума сигнала, энергии сигнала и максимума АЧХ для 4-х типов ЦФ с ЛОС.

Показано, что при масштабировании ЦФ с ЛОС 1-го и 2-го типов определяющую роль играют передаточные функции ЦФ1, ЦФ2, $H_{\text{ЛОС}2}(z)$. При масштабировании ЦФ с ЛОС 3-го и 4-го типов определяющую роль играют передаточные функции ЦФ1 и $H_{\text{ЛОС}2}(z)$. Определена дисперсия шума на выходе НЦФ с НЛОС после масштабирования для различных схем реализации. Показано, что с точки зрения помехоустойчивости наилучшей является типовая схема 3, а наихудшей – типовая схема 1. Причем помехоустойчивость ЦФ с ЛОС при нерекурсивной реализации будет выше, чем при рекурсивной реализации.

Таким образом, полученные результаты по масштабированию нерекурсивных ЦФ с нерекурсивными ЛОС позволяют, с одной стороны, предотвратить переполнение разрядной сетки, а с другой – обеспечить максимально возможное отношение сигнала к шуму округления по сравнению с рекурсивной реализацией.

Приведенные в статье данные по масштабированию ЦФ с ЛОС с использованием импульсных и передаточных функций для типовых одноуровневых двухканальных ЦФ с ЛОС можно обобщить и на n -канальную многоуровневую систему с различной структурой охвата каналов ЛОС.

Литература

1. Гольденберг Л. М., Матюшкин Б. Д., Поляк М. Н. Цифровая обработка сигналов. М.: Радио и связь, 1990. 256 с.
2. Рабинер Л., Голд Б. Теория и применение цифровой обработки сигналов. М.: Мир, 1978. 848 с.
3. Бакалов В. П., Субботин Е. А. Обработка сигналов в системах с латеральными связями. М.: Горячая линия – Телеком, 2017. 178 с.
4. Бакалов В. П., Черных Ю. С. Использование принципа латеральных связей для обработки дискретных сигналов // Вестник СибГУТИ. 2012. № 1. С. 43–50.
5. Бакалов В. П., Субботин Е. А. Исследование помехоустойчивости дискретных фильтров с латеральными связями // Вестник СибГУТИ. 2015. № 3. С. 89–95.
6. Бакалов В. П., Черных Ю. С. О потенциальной скорости передачи информации в инфокоммуникационных сетях с латеральными связями. // Вестник СибГУТИ. 2012. № 2. С. 62–70.
7. Бакалов В. П., Крук Б. И. Теория электрических цепей. М.: Горячая линия – Телеком, 2018. 537 с.
8. Бакалов В. П., Субботин Е. А. Масштабирование цифровых фильтров с латеральными связями // Вестник СибГУТИ. 2018. № 4. С. 12–19.

Статья поступила в редакцию 14.03.2019.

Бакалов Валерий Пантелеевич

д.т.н., профессор, заведующий кафедрой теории электрических цепей СибГУТИ (630102, Новосибирск, ул. Кирова, 86), тел. 286-80-35, e-mail: bvp@sibsutis.ru.

Субботин Евгений Андреевич

к.т.н., доцент, УрТИСИ СибГУТИ, тел. (343) 242-14-83.

Scaling nonrecursive of digital filters with nonrecursive lateral feedbacks

V. P. Bakalov, E. A. Subbotin

The article states ways to scale digital filters (DF) with lateral feedbacks (LFB). Models of four types of DF with LFBs with their nonrecursive implementation are considered. Filters are scaled using hi impulse response and scaling using the amplitude-frequency characteristic of the circuit. The obtained results allow, on the one hand, to prevent overfilling of the grade graticule of the DF with LFBs, and on the other hand, to significantly increase its noise immunity and efficiency.

Keywords: nonrecursive digital filters, nonrecursive lateral feedbacks, scaling, limiting the maximum signal, maximum signal energy, maximum amplitude-frequency characteristic of the circuit.

Математическое моделирование переходных термических процессов с учётом явлений тепловой релаксации и термического демпфирования

Л. С. Петрова, Е. В. Заец

В статье представлена математическая модель с двухфазным запаздыванием на основе уравнения теплопроводности гиперболического типа, учитывающего явления тепловой релаксации и термического демпфирования. Получено численное решение гиперболической задачи теплопроводности с учётом конечной скорости распространения тепла и демпфирования температуры. Рассмотрена реализация метода сеток с использованием трёхслойной неявной разностной схемы при решении задачи нестационарной теплопроводности для переходных термических процессов. Описан алгоритм расчёта температурного поля с использованием уравнения с двухфазным запаздыванием. Представлены результаты расчётов температурных полей в биологических тканях по уравнению теплопроводности гиперболического типа для расчётной области в пределах двухфазной зоны.

Ключевые слова: математическая модель, уравнение теплопроводности гиперболического типа, модель с двухфазным запаздыванием, численные методы, метод прогонки.

1. Введение

В настоящее время исследование быстропротекающих (высокоскоростных) переходных термических процессов является перспективным направлением в области интенсификации технологических процессов с достижением сверхвысоких и сверхнизких температур, использования лазерного облучения и материалов со сложной структурой.

В современных научных исследованиях (Ю. А. Кирсанов, А. Ю. Кирсанов, А. Е. Юдахин [1], Л. А. Фильштинский, Т. А. Киричек [2], Э. М. Карташов [3], Н. Z. Roog, Н. Moosavi, А. Moradi [4] и др.) отмечается ограниченность применения для быстропротекающих переходных процессов классической теории теплопроводности, основанной на гипотезе Фурье, содержащей допущение о бесконечной скорости распространения тепла, и приводится более сложная зависимость между тепловым потоком и градиентом температуры, которая учитывает тепловую релаксацию и термическое демпфирование [1, с. 70]:

$$q + \tau_q \frac{\partial q}{\partial \tau} = -\lambda \nabla T - \tau_T \lambda \frac{\partial \nabla T}{\partial \tau}, \quad (1)$$

где τ_q , τ_T – два несвязанных времени запаздывания (τ_q – время тепловой релаксации, τ_T – время термического демпфирования), q – плотность теплового потока, ∇T – градиент температуры, τ – время, λ – теплопроводность.

Сочетание формулы (1) и уравнения теплового баланса $c\rho \frac{\partial T}{\partial \tau} = -\operatorname{div} q$ позволило исследователям получить гиперболическое уравнение теплопроводности, учитывающее скорость изменения плотности потока тепла и скорость изменения градиента температуры [5, с. 550]:

$$\frac{\partial T}{\partial \tau} + \tau_q \frac{\partial^2 T}{\partial \tau^2} = a \Delta \left(T + \tau_T \frac{\partial T}{\partial \tau} \right), \quad (2)$$

где $a = \frac{\lambda}{c\rho}$ – коэффициент температуропроводности, $\Delta T = \nabla^2 T = \operatorname{div}(\operatorname{grad} T)$ – оператор Лапласа.

В исследованиях российских авторов для уравнения (1) используется название – уравнение Максвелла–Каттанео–Лыкова [6, с. 18] или уравнение с двухфазным запаздыванием [2, с. 164], иностранными исследователями это уравнение называется уравнением двухфазного запаздывания (dual-phase-lag) [5, с. 550].

В настоящее время исследователями рассматриваются математические модели с запаздыванием в различных прикладных областях науки и техники. Использование дифференциальных уравнений в частных производных с эффектом запаздывания при описании математических моделей динамических процессов в электродинамике, механике деформированного твердого тела, медицине, экологии и экономике отражено в работах Ю. Ф. Долгия, П. Г. Суркова [7], В. Г. Пименова, А. Б. Ложникова [8], А. И. Цаплина, И. Л. Никулина [9], Н. А. Быкова, О. К. Овчинниковой [10] и других.

Использование аналитических методов решения задач нестационарной теплопроводности с двухфазным запаздыванием на основе интегральных преобразований представлено в работах [1–5, 11–14]. Применение численных методов для решения гиперболической задачи теплопроводности с учётом конечной скорости распространения тепла и демпфирования температуры остается мало исследованным [9, 10]. В работе Н. А. Быкова и О. К. Овчинниковой [10, с. 144–146] реализуется метод ловли фронта в узел сетки с переменным шагом по времени (метод с выделением границы раздела фаз), при этом математическая постановка задачи сводится к системе двух уравнений теплопроводности параболического типа. Исследователями А. И. Цаплиным, И. Л. Никулиным [9, с. 117–121] используется метод сквозного счета, но с введением эффективной теплоемкости уравнение теплопроводности для расчетной области в пределах двухфазной зоны приводится к стандартному виду. Исследований по тематике применения численных методов с реализацией неявной разностной схемы при решении задач гиперболической теплопроводности с учётом явлений тепловой релаксации и термического демпфирования нами не обнаружено.

2. Задача нестационарной теплопроводности для переходных термических процессов в биологических тканях с учётом конечной скорости распространения тепла и демпфирования температуры

Значительный интерес у исследователей вызывает математическое моделирование быстропротекающих переходных термических процессов в биологических тканях. Реализация аналитических методов решения задач теплопроводности в условиях быстропротекающего термического процесса в материалах с низкой теплопроводностью представлена в исследованиях F. Filipoiu, A. I. Bogdan, I. M. Cârstea [11], H. Z. Poor, H. Moosavi, A. Moradi [4], Y. Zhang [12], K. C. Liu, P. J. Cheng, Y. N. Wang [13], E. A. N. Al-Lehaibi [14]. В работах [4] и [13] описано применение метода разделения переменных и операционного метода для расчёта температурного поля с использованием уравнения с двухфазным запаздыванием. Матема-

тическое моделирование на основе уравнения теплопроводности параболического типа для расчётной области в пределах двухфазной зоны представлено в исследованиях [11] и [12]. Использование операционного метода при решении задачи теплопроводности с двухфазным запаздыванием на основе перехода к системе двух уравнений теплопроводности (параболического и гиперболического типов) описано в работе [14]. Использование операционного метода при решении краевых задач с уравнением теплопроводности гиперболического типа, учитывающим явления тепловой релаксации и демпфирования температуры, в случае постоянного, периодического и импульсного тепловых потоков на границе сопряжено с возрастанием трудоёмкости выполняемых преобразований и актуализирует проблему использования численных методов решения.

Математическая модель быстропротекающего процесса теплопередачи в тканях кожи с постоянным тепловым потоком на поверхности включает в себя уравнение теплопроводности с двухфазным запаздыванием [4, с. 1460]:

$$\begin{aligned} \tau_q \rho c \frac{\partial^2 T(x, \tau)}{\partial \tau^2} + (\rho c + \tau_q \rho_b c_b \varpi_b) \frac{\partial T(x, \tau)}{\partial \tau} + \rho_b c_b \varpi_b (T(x, \tau) - T_a) = \\ = \lambda \left(\frac{\partial^2 T(x, \tau)}{\partial x^2} + \tau_T \frac{\partial^3 T(x, \tau)}{\partial x^2 \partial \tau} \right) + Q_{met} + Q_{ext} + \tau_q \frac{\partial Q_{met}}{\partial \tau} + \tau_q \frac{\partial Q_{ext}}{\partial \tau}, \\ 0 < x < L, \quad 0 < \tau \leq \tau_m; \end{aligned} \quad (3)$$

начальные условия:

$$T(x, 0) = T_a, \quad (4)$$

$$\left. \frac{\partial T(x, \tau)}{\partial \tau} \right|_{\tau=0} = 0, \quad x \in [0, L]; \quad (5)$$

условия на границах:

$$-\lambda \left. \frac{\partial T(x, \tau)}{\partial x} \right|_{x=0} = q_0, \quad (6)$$

$$\left. \frac{\partial T(x, \tau)}{\partial x} \right|_{x=L} = 0, \quad 0 \leq \tau \leq \tau_m, \quad (7)$$

где $T(x, \tau)$ – температура ткани в точке x в момент времени τ , λ – коэффициенты теплопроводности ткани, c – теплоемкость ткани, ρ – плотность ткани, c_b – теплоемкость крови, ρ_b – плотность крови, ϖ_b – скорость перфузии крови, T_a – температура крови, Q_{met} – метаболическое тепловыделение в ткани, Q_{ext} – мощность внешних тепловых источников, q_0 – плотность теплового потока. В этом исследовании $Q_{met} = \text{const}$, $Q_{ext} = 0$.

3. Метод сеток с использованием трёхслойной неявной разностной схемы при решении гиперболической задачи теплопроводности для расчётной области в пределах двухфазной зоны

Реализуя метод сеток для решения задачи (3) – (7), вводим прямоугольную пространственно-временную сетку $G_{x\tau} = \{x_i = (i-1)h, i = \overline{1, N}; \tau_j = jk, j = \overline{0, M}\}$. Для получения трёхслойной неявной разностной схемы используем формулы соответствующих разностных ана-

логов частных производных первого и второго порядков по времени, производной второго порядка по пространству, представленные в работе [15, с. 33–36]. Применение теории разностных схем, описанной в работе [15, с. 71–80], позволило получить разностный аналог для смешанной производной третьего порядка с погрешностью аппроксимации $O(kh^2)$:

$$\frac{\partial^3 T(x, \tau)}{\partial x^2 \partial \tau} \approx \frac{T_{i+1}^{j+1} - T_{i+1}^j - 2(T_i^{j+1} - T_i^j) + T_{i-1}^{j+1} - T_{i-1}^j}{h^2 k}. \quad (8)$$

Разностное уравнение в этом случае имеет вид:

$$\begin{aligned} \tau_q \rho c \frac{T_i^{j+1} - 2T_i^j + T_i^{j-1}}{k^2} + (\rho c + \tau_q \rho_b c_b \varpi_b) \frac{T_i^{j+1} - T_i^j}{k} + \rho_b c_b \varpi_b (T_i^{j+1} - T_a) = \\ = \lambda \left(\frac{T_{i+1}^{j+1} - 2T_i^{j+1} + T_{i-1}^{j+1}}{h^2} + \tau_T \frac{T_{i+1}^{j+1} - T_{i+1}^j - 2(T_i^{j+1} - T_i^j) + T_{i-1}^{j+1} - T_{i-1}^j}{h^2 k} \right) + Q_{met}, \end{aligned} \quad (9)$$

где $i = \overline{2, N-1}$, $j = \overline{1, M-1}$.

Модифицируя метод прогонки для семиточечной разностной схемы, систему (9) приводим к системе линейных алгебраических уравнений, имеющей трёхдиагональную структуру:

$$A_i \cdot T_{i+1}^{j+1} - B_i \cdot T_i^{j+1} + C_i \cdot T_{i-1}^{j+1} = F_i, \quad (10)$$

где $A_i = C_i = \lambda(k^2 + k\tau_T)$, $B_i = h^2(\tau_q + k)(\rho c + k\rho_b c_b \varpi_b) + 2\lambda k(k + \tau_T)$, $F_i = h^2 \tau_q \rho c T_i^{j-1} + \lambda \tau_T k (T_{i-1}^j + T_{i+1}^j) - T_i^j [2\lambda \tau_T k + h^2 \tau_q \rho c + kh^2(\rho c + \tau_q \rho_b c_b \varpi_b)] - k^2 h^2 (\rho_b c_b \varpi_b T_a + Q_{met})$.

Используя основное соотношение прогонки $T_{i-1}^{j+1} = \alpha_{i-1} \cdot T_i^{j+1} + \beta_{i-1}$ для уравнения (10), получаем формулы для расчёта прогоночных коэффициентов:

$$\alpha_i = \frac{A_i}{B_i - C_i \cdot \alpha_{i-1}}, \quad \beta_i = \frac{C_i \cdot \beta_{i-1} - F_i}{B_i - C_i \cdot \alpha_{i-1}}. \quad (11)$$

Определяя коэффициенты F_i на первом временном слое, используем соотношение для фиктивных узлов $T_i^{-1} = T_i^0$, полученное из аппроксимации второго начального условия:

$$F_i^1 = \lambda \tau_T k (T_{i-1}^0 + T_{i+1}^0) - T_i^0 [2\lambda \tau_T k + h^2 \tau_q \rho c + kh^2(\rho c + \tau_q \rho_b c_b \varpi_b)] - k^2 h^2 (\rho_b c_b \varpi_b T_a + Q_{met}) \quad (12)$$

Для получения начальных прогоночных коэффициентов α_1, β_1 используем разностное уравнение (9) и соотношение для значений сеточной функции в фиктивных узлах, полученное из аппроксимации левого граничного условия с погрешностью $O(h^2)$:

$$\alpha_1^{j+1} = \frac{2\lambda k(k + \tau_T)}{h^2(\tau_q + k)(\rho c + k\rho_b c_b \varpi_b) + 2\lambda k(k + \tau_T)}, \quad (13)$$

$$\begin{aligned} \beta_1^{j+1} = \frac{2hk^2 q_0 + 2\lambda \tau_T k (T_1^j - T_2^j) + h^2 \rho c [T_1^j (2\tau_q + k) - T_1^{j-1} \tau_q]}{h^2(\tau_q + k)(\rho c + k\rho_b c_b \varpi_b) + 2\lambda k(k + \tau_T)} + \\ + \frac{kh^2 \rho_b c_b \varpi_b (T_1^j \tau_q + kT_a) + Q_m k^2 h^2}{h^2(\tau_q + k)(\rho c + k\rho_b c_b \varpi_b) + 2\lambda k(k + \tau_T)}. \end{aligned} \quad (14)$$

Начальные прогоночные коэффициенты на первом временном слое определяем, используя соотношение для фиктивного узла $T_1^{-1} = T_1^0$:

$$\beta_1^1 = \frac{2\lambda\tau_T k(T_1^0 - T_2^0) + h^2 \rho c T_1^0 (\tau_q + k) + kh^2 \rho_b c_b \varpi_b (T_1^0 \tau_q + kT_a) + hk^2 (2q_0 + Q_m h)}{h^2 (\tau_q + k)(\rho c + k\rho_b c_b \varpi_b) + 2\lambda k(k + \tau_T)}. \quad (15)$$

Формулы для вычисления значений температуры на правой границе получаем, используя аппроксимацию правого граничного условия с погрешностью $O(h^2)$ и значение сеточной функции в фиктивном узле $T_N^{-1} = T_N^0$ на первом временном слое:

$$T_N^{j+1} = \frac{h^2 \rho c (\tau_q T_N^{j-1} - 2\tau_q T_N^j - kT_N^j) - kh^2 \rho_b c_b \varpi_b (\tau_q T_N^j + kT_a)}{2\lambda k(\tau_T + k)(\alpha_{N-1} - 1) - h^2 (\tau_q + k)(\rho c + k\rho_b c_b \varpi_b)} +$$

$$+ \frac{2\lambda\tau_T k(T_{N-1}^j - T_N^j) - Q_m k^2 h^2 - 2\lambda k \beta_{N-1}(\tau_T + k)}{2\lambda k(\tau_T + k)(\alpha_{N-1} - 1) - h^2 (\tau_q + k)(\rho c + k\rho_b c_b \varpi_b)}; \quad (16)$$

$$T_N^1 = \frac{-h^2 \rho c T_N^0 (\tau_q + k) - kh^2 \rho_b c_b \varpi_b (\tau_q T_N^0 + kT_a)}{2\lambda k(\tau_T + k)(\alpha_{N-1} - 1) - h^2 (\tau_q + k)(\rho c + k\rho_b c_b \varpi_b)} +$$

$$+ \frac{2\lambda\tau_T k(T_{N-1}^0 - T_N^0) - Q_m k^2 h^2 - 2\lambda k \beta_{N-1}(\tau_T + k)}{2\lambda k(\tau_T + k)(\alpha_{N-1} - 1) - h^2 (\tau_q + k)(\rho c + k\rho_b c_b \varpi_b)}. \quad (17)$$

При реализации модификации метода прогонки, предусматривающей использование процедуры расчёта значений сеточной функции на промежуточном и первом временных слоях, на этапе прямой прогонки для определения прогоночных коэффициентов используются формулы (11), (13) – (15). Расчёт температурного поля на этапе обратной прогонки осуществляется с применением формул (16) и (17) для вычисления значений температуры на правой границе, а также с применением основного соотношения прогонки для остальных узлов на текущем временном слое.

Отличие знаменателей прогоночных коэффициентов от нуля и выполнение условия $|\alpha_i| < 1, i = \overline{1, N-1}$, исключающего быстрый рост погрешности округления, обеспечивает успешное применение метода прогонки с учётом безусловной устойчивости предложенной неявной разностной схемы с погрешностью аппроксимации $O(k + h^2)$.

4. Результаты расчётов температурных полей с использованием уравнения с двухфазным запаздыванием

Реализация программы численного решения задачи (3) – (7) осуществлялась в системе MathCAD, в среде программирования Visual Studio Express Edition на языке C# и в среде Dev-C++.

Сравнение и анализ результатов расчётов температурных полей на основе предложенного численного решения проводились при следующих значениях теплофизических характеристик и геометрических условиях [4, с. 1465, 13, с. 65]: $\lambda = 0,235 \text{ Вт}/(\text{м} \cdot \text{К})$, $\rho = 1190 \text{ кг}/\text{м}^3$,

$\rho_b = 1060 \text{ кг/м}^3$, $c = 3600 \text{ Дж/(кг} \cdot \text{К)}$, $c_b = 3770 \text{ Дж/(кг} \cdot \text{К)}$, $W_b = \rho_b \varpi_b = 0,5 \text{ кг/(м}^3\text{с)}$,
 $T_a = 37 \text{ }^\circ\text{C}$, $Q_{met} = 368,1 \text{ Вт/м}^3$, $q_1 = 5000 \text{ Вт/м}^2$, $q_2 = 2000 \text{ Вт/м}^2$, $L = 6 \text{ мм}$, $\tau_T = 10 \text{ с}$,
 $\tau_q = 16 \text{ с}$.

Результаты расчётов температурных полей с постоянным тепловым потоком на поверхности $q_1 = 5000 \text{ Вт/м}^2$ (T_1), $q_2 = 2000 \text{ Вт/м}^2$ (T_2) представлены в виде графиков распределения температуры по толщине слоя при $\tau = 15 \text{ с}$, $\tau = 45 \text{ с}$ и функции изменения температуры при $x = 0,1 \text{ мм}$, $x = 1,6 \text{ мм}$ на рис. 1 – 2.

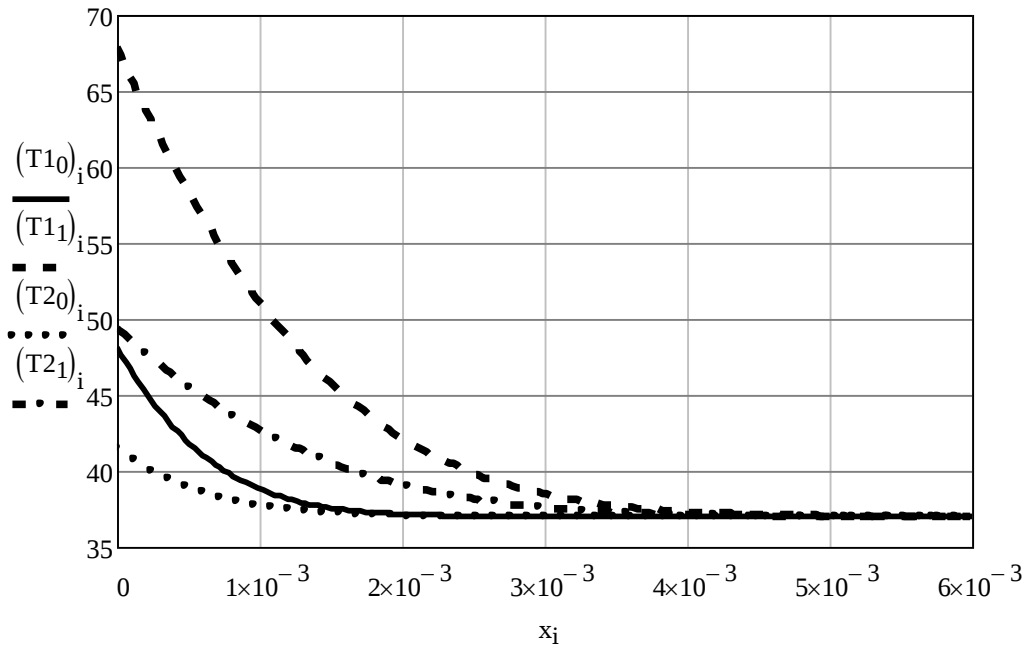


Рис. 1. Распределение температуры по толщине слоя: при $\tau = 15 \text{ с}$ (T_{10} , T_{20}) и при $\tau = 45 \text{ с}$ (T_{11} , T_{21})

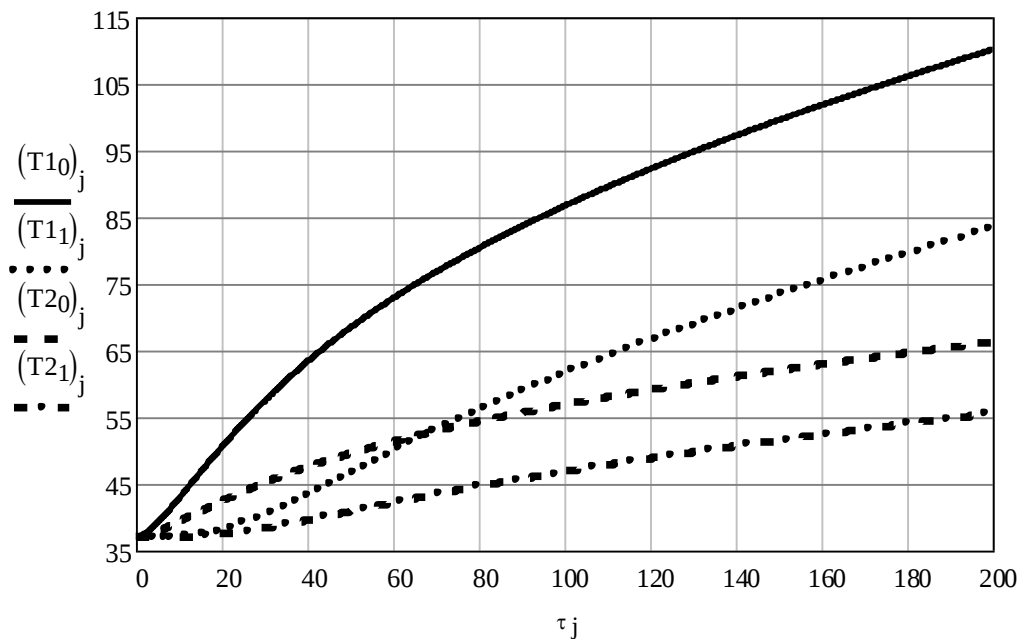


Рис. 2. Изменение температуры во времени: при $x = 0,1 \text{ мм}$ (T_{10} , T_{20}) и при $x = 1,6 \text{ мм}$ (T_{11} , T_{21})

Сопоставимость представленных результатов с результатами расчётов, полученными в работе [4, с. 1469–1471] с применением аналитических методов на основе метода разделения переменных, обосновывает достоверность полученных результатов.

5. Заключение

Исследование явлений теплопередачи в условиях быстропротекающего термического процесса в материалах с низкой теплопроводностью (например, в биологических тканях) имеет большое значение и может найти разнообразное применение, в том числе для разработки медицинских технологий, конструирования оборудования и многих других целей.

Разработанная математическая модель с двухфазным запаздыванием может использоваться при анализе технологических процессов, основанных на теплопередаче быстропротекающих переходных термических процессов в тканях кожи с постоянным тепловым потоком на поверхности, для расчётов температурных полей в биологических тканях с использованием уравнения теплопроводности гиперболического типа, учитывающего явления тепловой релаксации и термического демпфирования.

Литература

1. Кирсанов Ю. А., Кирсанов А. Ю., Юдахин А. Е. Решение краевой гиперболической задачи теплопроводности для переходного термического процесса // Материалы X школы-семинара молодых учёных и специалистов академика РАН В. Е. Алемасова «Проблемы тепломассообмена и гидродинамики в энергомашиностроении». 2016. С. 70–73.
2. Фильштинский Л. А., Киричек Т. А. Применение неклассических моделей теплопроводности для расчётов тепловых полей в элементах конструкций // Авиационно-космическая техника и технология. 2005. № 7 (23). С. 162–170.
3. Карташов Э. М. Математические модели теплопроводности с двухфазным запаздыванием // Инженерно-физический журнал. 2016. № 2 (89). С. 338–349.
4. Poor H. Z., Moosavi H., Moradi A. Analysis of the DPL bio-heat transfer equation with constant and time-dependent heat flux conditions on skin surface // Thermal Science. 2016. V. 20, № 5. P. 1457–1472.
5. Кирсанов Ю. А. Влияние тепловой релаксации и термического демпфирования на переходные процессы при циклических граничных условиях // Теплофизика высоких температур. 2017. № 4 (55). С. 549–555.
6. Юдахин А. Е. Исследование моделей теплопроводности в условиях быстропротекающего термического процесса в низкотеплопроводном твердом теле: дис. ... канд. техн. наук. Казань. 2017. 141 с.
7. Долгий Ю. Ф., Сурков П. Г. Математические модели динамических систем с запаздыванием: учебное пособие. Екатеринбург: Изд-во Урал. ун-та, 2012. 122 с.
8. Пименов В. Г., Ложников А. Б. Численный метод моделирования управляемого уравнения теплопроводности с запаздыванием // Вестник Тамбовского университета. Серия. Естественные и технические науки. 2013. № 5-2 (18). С. 2635–2636.
9. Цаплин А. И., Никулин И. Л. Моделирование теплофизических процессов и объектов в металлургии: учебное пособие. Пермь: Изд-во Перм. гос. техн. ун-та, 2011. 299 с.
10. Брыков Н. А., Овчинникова О. К. Численное решение задачи плавления твердого вещества // Успехи современной науки. 2017. № 3 (6). С. 144–147.
11. Filipoiu F., Bogdan A. I., Cârstea I. M. Computer-aided analysis of the heat transfer in skin tissue // Proceedings of the 3rd WSEAS Int. Conference on finite differences – finite elements – finite volumes – boundary elements. С. 53–59.

12. Zhang Y. Generalized dual-phase lag bioheat equations based on nonequilibrium heat transfer in living biological tissues // International Journal of Heat and Mass Transfer. 2009. V. 52. P. 4829–4834.
13. Liu K. C., Cheng P. J., Wang, Y. N. Analysis of non-Fourier thermal behavior for multi-layer skin model // Thermal Science. 2011. V. 15 (1). P. 61–67.
14. Al-Lehaibi E. A. N. The skin tissue of the human head subjected to thermal diffusion // Mathematical Problems in Engineering. 2018. [Электронный ресурс]
URL: <https://doi.org/10.1155/2018/8781950> (дата обращения 08.02.2019).
15. Самарский А. А. Теория разностных схем: учебное пособие для вузов. М.: Наука, 1977. 656 с.

Статья поступила в редакцию 19.02.2019.

Петрова Лилия Сергеевна

к.пед.н., доцент по специальности «Математическое моделирование, численные методы и комплексы программ», доцент кафедры «Высшая математика», Омский государственный университет путей сообщения (ОмГУПС) (644046, Омск, пр. Маркса, 35), e-mail: petrov.306@mail.ru.

Заец Евгений Валерьевич

магистрант кафедры «Теплоэнергетика», Омский государственный университет путей сообщения, e-mail: zaetsevgenijmagistr@gmail.com.

Mathematical modeling of transition thermal processes taking into account effect of thermal relaxation and thermal damping.

L. S. Petrova, E. V. Zaets

The article presents a mathematical model with dual phase lag based on the heat conduction equation of hyperbolic type, taking into account effect of thermal relaxation and thermal damping. The numerical solution of the hyperbolic heat conduction problem is obtained taking into account the final speed of heat spread and temperature damping. The implementation of the grid method using a three-layer implicit difference scheme in solving the problem of non-stationary heat conductivity for a transition thermal processes were treated. An algorithm for calculating the temperature field using a dual phase lag equation is described. The results of calculations of temperature fields in biological tissues according to the hyperbolic thermal conductivity equation for the calculation area within the dual phase zone are presented.

Keywords: mathematical model, heat conduction equation of hyperbolic type, model with dual phase lag, numerical methods, sweep method.

Измерение разности фаз гармонических сигналов в линейных трактах с квазибелым шумом

В. Г. Арсентьев, Г. И. Криволапов

Рассмотрен квадратурный метод измерения разности фаз гармонических сигналов с использованием преобразования частоты. Предложено ввести в алгоритмическую структуру метода процедуру синфазирования сигналов. Получена оценка эффективности разработанного модифицированного квадратурного метода.

Ключевые слова: разность фаз, квадратурный метод, синфазирование сигналов.

1. Введение

Задача повышения точности измерений разности фаз гармонических сигналов в линейных трактах различного назначения, имеющих ограниченные полосы частот с квазибелыми шумами, была и по-прежнему остаётся актуальной. Востребованность выполнения точных измерений имеет место, например, в локационных и навигационных системах, при измерениях фазочастотных характеристик трактов различных радиоэлектронных устройств, неразрушающем контроле и измерениях свойств материалов, а также в других прикладных областях, использующих фазовые измерения.

Для измерения разности фаз гармонических сигналов разработано достаточно много методов, например, таких как преобразование интервала времени в напряжение [1], дискретного счёта [2–4], измерения разности фаз с преобразованием частоты [1, 2], корреляционного измерения [1, 2, 5], преобразования Фурье [2, 6–8], наименьших квадратов [9].

Используемые на практике методы имеют разные точности измерения разности фаз, зависящие от функциональных особенностей самого метода и условий его применения, но в подавляющем большинстве для обеспечения высокой точности измерения (на уровне десятых долей градуса) требуется большое (более 40 дБ) отношение сигнал/квазибелый шум в исследуемых линейных трактах [10].

Наибольшее распространение в широкодиапазонных измерителях разности фаз получил метод измерения разности фаз с преобразованием частоты, который позволяет получить достаточно высокую точность измерения в условиях сравнительно небольших (около 30 дБ) отношений сигнал/квазибелый шум в исследуемых трактах за счёт гетеродинирования измерительных сигналов на низкую промежуточную частоту, но не обеспечивает требуемой точности измерения при меньших отношениях.

В качестве примера на рис. 1 показан возможный состав обобщённых функциональных элементов, необходимых в процессе измерения разности фаз сигналов с использованием преобразования частоты. Формируемый гармонический измерительный сигнал $u_n(t)$ с частотой f_n поступает на входы двух исследуемых трактов. Исследуемые тракты представляют собой устройства или физические среды, на выходе которых наблюдаются измерительные сигналы $x(t)$, $y(t)$, разность фаз которых $\Delta\varphi_n$ подлежит измерению. Тракты преобразования частоты включают в себя устройства (смесители, полосовые усилители и общий гетеродин с частотой

$f_r = f_n - f_{пр}$), осуществляющие перенос выходных измерительных сигналов исследуемых трактов на низкую промежуточную частоту $f_{пр}$ для выполнения дальнейших измерений разности фаз сигналов $X(t)$, $Y(t)$ на базе одного из известных методов.

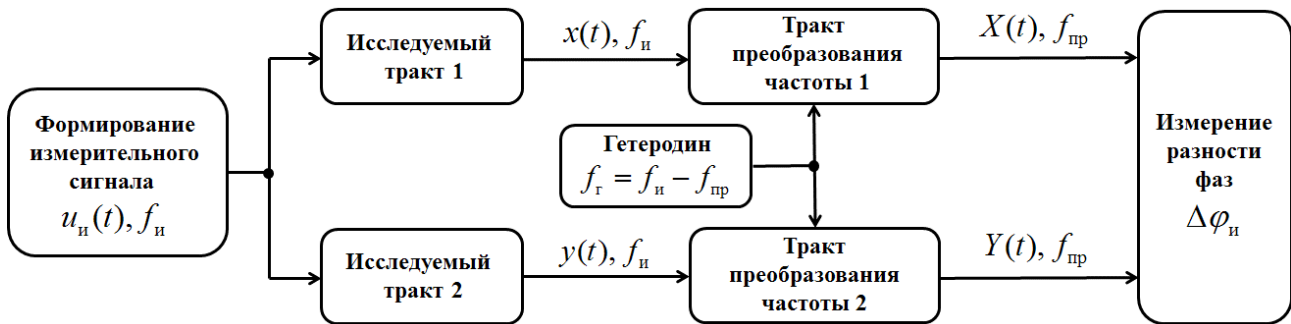


Рис. 1. Измерение разности фаз сигналов двух исследуемых трактов

В статье рассматривается разработанный авторами модифицированный квадратурный метод измерения разности фаз на низкой промежуточной частоте, использующий процедуру синфазирования сигналов, позволяющую повысить точность измерений в условиях малых отношений сигнал/квазибелый шум исследуемых трактов.

Оценка эффективности предлагаемого метода выполнена с использованием среды математического моделирования Mathcad 15.

2. Квадратурный метод измерения разности фаз сигналов

Измерительный гармонический сигнал $u_n(t)$, поступающий на входы исследуемых трактов, имеет вид

$$u_n(t) = U_n \cos(2\pi f_n t + \varphi_n), \quad (1)$$

где U_n , f_n , φ_n – амплитуда, частота и начальная фаза колебаний соответственно.

Измерительные сигналы $x(t)$, $y(t)$ на выходах исследуемых трактов характеризуются соотношениями:

$$\begin{aligned} x(t) &= U_1 \cos(2\pi f_n t + \varphi_n + \varphi_1) + \xi_1(t); \\ y(t) &= U_2 \cos(2\pi f_n t + \varphi_n + \varphi_2) + \xi_2(t); \\ \Delta\varphi_n &= \varphi_1 - \varphi_2; \quad \Delta\varphi_n \in [-\pi \div +\pi], \end{aligned} \quad (2)$$

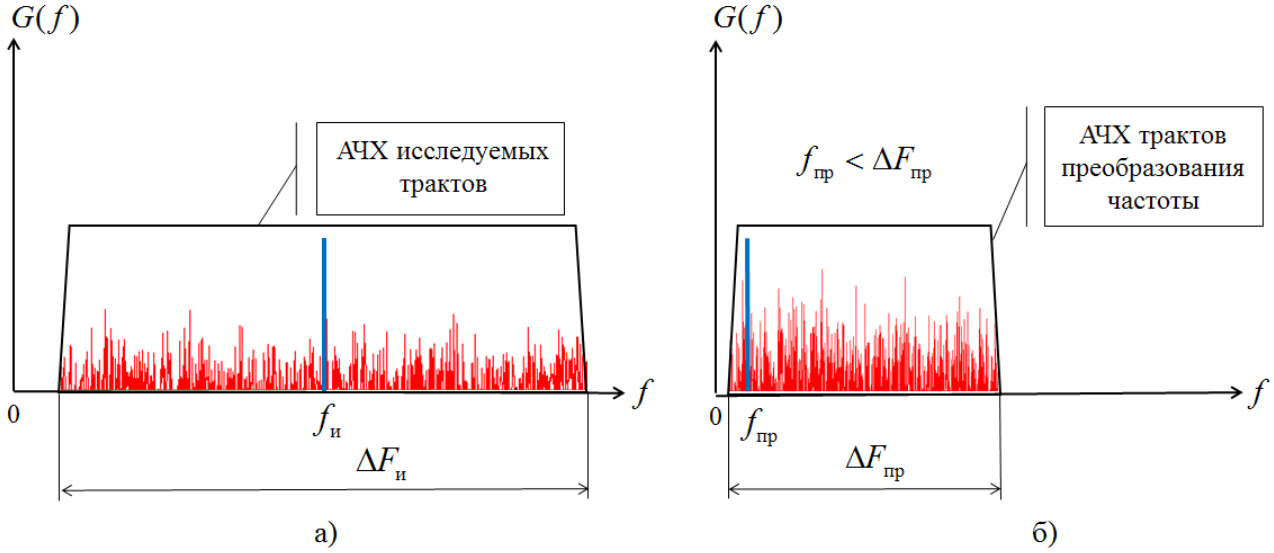
где U_1 , U_2 – амплитуды колебаний; φ_1 , φ_2 – вносимые исследуемыми трактами фазовые сдвиги колебаний; $\Delta\varphi_n$ – измеряемая разность фаз колебаний; $\xi_1(t)$, $\xi_2(t)$ – реализации квазибелых шумов исследуемых трактов.

Спектры измерительного сигнала и квазибелого шума на выходах исследуемых трактов с полосами частот ΔF_n показаны на рис. 2а.

Измерительные сигналы $x(t)$, $y(t)$ исследуемых трактов переносятся на промежуточную частоту $f_{пр}$ с использованием общего гетеродина, колебания которого $u_r(t)$ имеют следующий вид

$$u_r(t) = U_r \cos [2\pi(f_n - f_{пр})t + \varphi_r], \quad (3)$$

где U_r , φ_r – амплитуда и начальная фаза колебаний; $f_{пр}$ – промежуточная частота преобразования.



АЧХ – амплитудно-частотные характеристики

Рис. 2. Спектры измерительного сигнала и квазибелого шума на выходах исследуемых трактов и трактов преобразования частоты

Выходные сигналы $X(t)$, $Y(t)$ исследуемых трактов после преобразования частоты и фильтрации колебаний разностной частоты описываются следующими зависимостями:

$$\begin{aligned} X(t) &= U_x \cos(2\pi f_{пр} t + \varphi_n + \varphi_1 - \varphi_r) + \xi_3(t); \\ Y(t) &= U_y \cdot \cos(2\pi f_{пр} t + \varphi_n + \varphi_2 - \varphi_r) + \xi_4(t), \end{aligned} \quad (4)$$

где U_x , U_y – амплитуды колебаний измерительных сигналов на выходах трактов преобразования частоты (с учётом коэффициентов передачи смесителей и полосовых усилителей); $\xi_3(t)$, $\xi_4(t)$ – реализации квазибелых шумов преобразовательных трактов.

Спектры измерительного сигнала и квазибелого шума на выходах трактов преобразования частоты с полосами частот $\Delta F_{пр}$ приведены на рис. 2б, где для корректности дальнейших сравнительных расчётов обеспечено выполнение условия: $f_{пр} < \Delta F_{пр}$.

Если сформировать из сигнала $X(t)$ синфазную и квадратурную составляющие $X_c(t)$ и $X_s(t)$ с амплитудами U_x (техническая реализация этой операции может быть самой разнообразной) и осуществить усреднение на интервале времени, значительно превышающем их период, произведений двух пар сигналов $X_c(t) \cdot Y(t)$ и $X_s(t) \cdot Y(t)$, то получим

$$\Lambda_1 = \langle X_c(t) \cdot Y(t) \rangle \approx \frac{1}{2} U_x U_y \cos \Delta \varphi_n; \quad (5)$$

$$\Lambda_2 = \langle X_s(t) \cdot Y(t) \rangle \approx \frac{1}{2} U_x U_y \sin \Delta \varphi_n, \quad (6)$$

где $\langle X \cdot Y \rangle$ знак операции усреднения по времени.

Амплитуды сигналов U_x , U_y находятся в линейной связи с усреднёнными модулями гармонических колебаний $X(t)$ и $Y(t)$:

$$\langle |X(t)| \rangle = \frac{1}{\Delta T} \int_0^{\Delta T} |X(t)| dt \approx \frac{2U_x}{\pi}; \quad \langle |Y(t)| \rangle = \frac{1}{\Delta T} \int_0^{\Delta T} |Y(t)| dt \approx \frac{2U_y}{\pi}; \quad \Delta T \gg \frac{1}{f_{\text{пр}}}. \quad (7)$$

Искомая разность фаз $\Delta\varphi_{\text{и}} = \varphi_1 - \varphi_2$ сигналов определяется из формулы (5) и соотношений (7) с учётом знака корреляционного коэффициента Λ_2 в выражении (6), приобретая вид

$$\Delta\varphi_{\text{и}} = \begin{cases} \arccos\left(\frac{2\Lambda_1}{U_x U_y}\right), & \text{если } \Lambda_2 \geq 0 \\ -\arccos\left(\frac{2\Lambda_1}{U_x U_y}\right), & \text{если } \Lambda_2 < 0 \end{cases} \quad [\text{рад.}]; \quad \Delta\varphi_{\text{и}} \in [-\pi \div +\pi]. \quad (8)$$

С точки зрения упрощения технической реализации и уменьшения количества операций квадратурного метода вычисления разности фаз представляет интерес использование предельного двухстороннего ограничения сигналов $X(t)$, $Y(t)$ после преобразования частоты.

С позиции математического описания предельным ограничителем называют нелинейное устройство, осуществляющее преобразование [11] вида

$$f(x) = A \cdot \text{sgn}(x) = A(2\chi(x) - 1), \quad (9)$$

где $\chi(x)$ – ступенчатая функция Хевисайда; A – действительная, не равная нулю константа.

При таком преобразовании сохраняется только знак исходного сигнала. Реализуется эта операция в случае, если сигнал сначала подвергается очень большому усилению, а затем ограничивается по амплитуде величиной A . Это преобразование известно как операция клиппирования сигнала, которая широко применяется в различных алгоритмах приёма дискретных сообщений и цифровой обработки сигналов.

На основе зависимости (9) можно установить функциональную связь корреляционного коэффициента Λ_{xy} двух клиппированных гармонических сигналов $X(t)$, $Y(t)$ и разностей фаз $\Delta\varphi_{\text{и}}$ их колебаний. Указанная связь описывается соотношением:

$$\Lambda_{xy} = \frac{2A^2}{\pi} \arcsin(\cos \Delta\varphi_{\text{и}}) = \frac{2A^2}{\pi} \left(\frac{\pi}{2} - \Delta\varphi_{\text{и}} \right), \quad (10)$$

где A – амплитуда каждого клиппированного сигнала.

Отсюда с учётом (8) несложно получить формулу для определения разности фаз квадратурного метода измерения с клиппированием сигналов после преобразования частоты:

$$\Delta\varphi_{\text{и}} = \begin{cases} \frac{\pi}{2} \left(1 - \frac{\Lambda_1}{A^2} \right), & \text{если } \Lambda_2 \geq 0 \\ -\frac{\pi}{2} \left(1 - \frac{\Lambda_1}{A^2} \right), & \text{если } \Lambda_2 < 0 \end{cases} \quad [\text{рад.}]; \quad \Delta\varphi_{\text{и}} \in [-\pi \div +\pi]. \quad (11)$$

При низкой промежуточной частоте преобразования ($f_{\text{пр}} \leq 10 \div 20$ кГц) и современной элементной базе целесообразна цифровая реализация квадратурного метода измерения разности фаз, которая предполагает следующие действия и операции.

Сигналы промежуточной частоты $X(t)$, $Y(t)$ подвергаются синхронной дискретизации с большим и кратным четырём количеством отсчётов k на периоде колебаний и общим числом отсчётов L в выборке, формируются массивы отсчётов $X(n)$, $Y(n)$:

$$X(t) \rightarrow X(n); \quad Y(t) \rightarrow Y(n); \quad n \in [1 \div L], \quad (12)$$

где n – порядковый номер отсчёта в выборке массива; $L = k \cdot (m + 0.25)$; m – целое число.

Для обеспечения малой погрешности вычислений разности фаз можно рекомендовать следующие значения параметров дискретизации и усреднения сигналов: $k = 3.6 \cdot 10^3$, $m = 100$.

Осуществляется нормирование отсчётов массивов $X(n)$, $Y(n)$ путём их представления своими знаками по правилу:

$$X(n) = \begin{cases} 1, & \text{если } X(n) > 0 \\ -1, & \text{если } X(n) < 0 \end{cases}; \quad Y(n) = \begin{cases} 1, & \text{если } Y(n) > 0 \\ -1, & \text{если } Y(n) < 0 \end{cases}; \quad n \in [1 \div L]. \quad (13)$$

Вычисляются корреляционные коэффициенты β и γ :

$$\beta = \frac{1}{L - 0.25k} \sum_{g=1}^{L-0.25k} X(g)Y(g); \quad \gamma = \frac{1}{L - 0.25k} \cdot \sum_{p=1+0.25k}^L X(p)Y(p-0.25k). \quad (14)$$

Рассчитывается искомая разность фаз $\Delta\varphi_{\text{и}}$ с использованием одной из предпочтительных формул:

$$\Delta\varphi_{\text{и}} = \begin{cases} \frac{\pi}{2}(1-\beta), & \text{если } \gamma \geq 0 \\ -\frac{\pi}{2}(1-\beta), & \text{если } \gamma < 0 \end{cases} [\text{рад.}]; \quad \Delta\varphi_{\text{и}} = \begin{cases} 90(1-\beta), & \text{если } \gamma \geq 0 \\ -90(1-\beta), & \text{если } \gamma < 0 \end{cases} [\text{град.}]; \quad (15)$$

$$\Delta\varphi_{\text{и}} \in [-\pi \div +\pi]; \quad \Delta\varphi_{\text{и}} \in [-180^\circ \div +180^\circ].$$

Следует заметить, что техническая реализация квадратурного метода измерения разности фаз с использованием преобразования частоты требует для точных измерений выполнения калибровки преобразовательных трактов перед каждым измерением на выбранной частоте измерительного сигнала.

Указанная процедура может быть осуществлена путём подачи в оба тракта преобразования частоты сигнала одного из исследуемых трактов ($x(t)$ или $y(t)$) и выполнения описанных выше действий и операций с получением разности фаз $\Delta\varphi_{\text{и}}^*$, которая будет являться калибровочной поправкой, вычитаемой затем из полученной далее при той же последовательности действий и операций разности фаз $\Delta\varphi_{\text{и}}$, но уже при подаче в тракты преобразования частоты обоих сигналов $x(t)$ и $y(t)$.

На рис. 3 в качестве примера представлены зависимости модуля абсолютной погрешности измерения разности фаз рассмотренного цифрового квадратурного метода от величины изме-

ряемой разности фаз для реализации квазибелого нормального шума с нулевым средним значением при коэффициенте широкополосности трактов преобразования частоты $N = \frac{\Delta F_{\text{пр}}}{f_{\text{пр}}} = 5$ и отношениях сигнал/квазибелый шум в указанных трактах $H = 20, 40$ дБ.

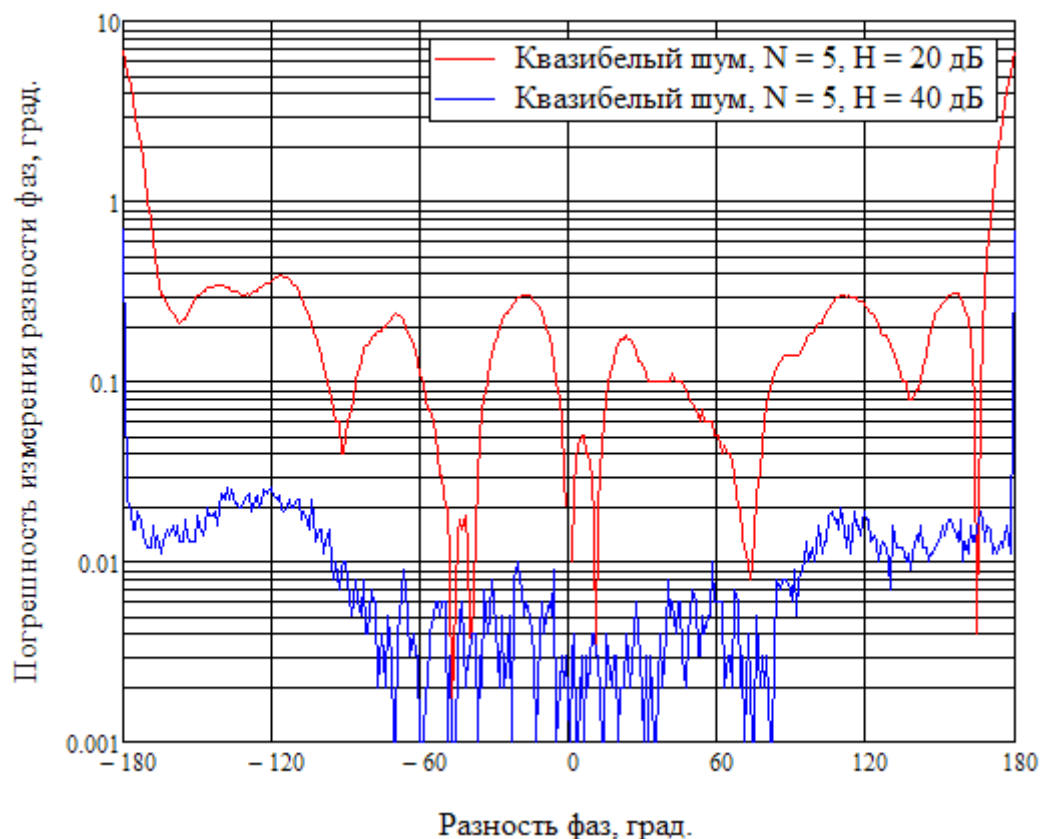


Рис. 3. Зависимость погрешности измерения от измеряемой разности фаз

Результаты моделирования указывают на достаточно высокую точность измерений цифрового квадратурного метода, выявляя характерные для него особенности:

- увеличение разброса погрешности измерения разности фаз с уменьшением отношения сигнал/квазибелый шум;
- значительное увеличение погрешности измерения разности фаз (максимальной погрешности метода) в областях значений, приближающихся к $|180^\circ|$, с расширением этих областей при уменьшении отношения сигнал/квазибелый шум.

Следует отметить, что указанная максимальная погрешность метода в диапазоне $H \in [20 \div 50]$ дБ уменьшается в три раза на каждые 10 дБ прироста H .

Указанные особенности можно отнести к недостаткам метода, однако они не являются принципиально ограничивающими факторами его сравнительно широкого применения.

3. Процедура синфазирования сигналов

Для осуществления процедуры формируется вспомогательный гармонический сигнал $Z(t)$ промежуточной частоты $f_{\text{пр}}$, имеющий большое (не менее 80 дБ) отношение сигнал/шум. В процедуре синфазирования добиваются синфазности колебаний в каждой паре сигналов промежуточной частоты: вспомогательного сигнала $Z(t)$ и сигнала $X(t)$ первого тракта преобразования частоты, вспомогательного сигнала $Z(t)$ и сигнала $Y(t)$ второго тракта преобразования частоты, что позволяет перейти от измерения разности фаз колебаний пары сигналов

трактов преобразования частоты с изменяющимися малыми отношениями сигнал/шум к измерению разности фаз колебаний пары вспомогательных сигналов, имеющих постоянно большее отношение сигнал/шум.

Погрешность измерения разности фаз зависит от точности синфазирования вышеупомянутых сигналов, которая достигается выбором величины полосы пропускания $\Delta F_{\text{пр}}$ трактов преобразования частоты, большей значения промежуточной частоты $f_{\text{пр}}$ (коэффициент широкополосности $N > 1$), при этом корреляционная обработка двух пар сигналов промежуточной частоты на фоне квазибелых шумов позволяет реализовать эффект когерентного накопления, хорошо известный в теории помехоустойчивой обработки сигналов [12].

Процедура синфазирования двух гармонических сигналов $X(t)$ и $Z(t)$ иллюстрируется рис. 4. Из вспомогательного сигнала $Z(t)$ формируются квадратурные составляющие $Z1(t)$ и $Z2(t)$, которые подвергаются весовой обработке с коэффициентами $W1$ и $W2$, минимизирующими разность $\varepsilon(t)$ сигналов $X(t)$ и взвешенного сигнала $Z(t)$.

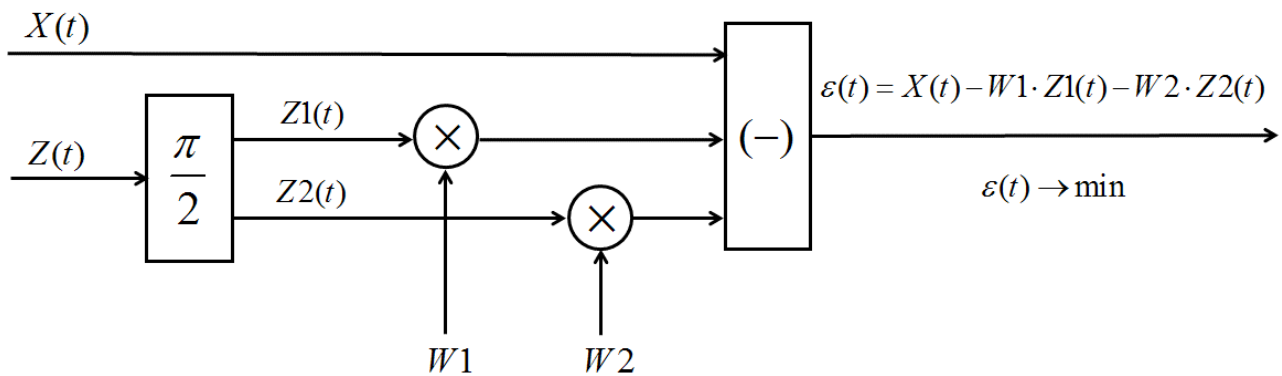


Рис. 4. Принцип синфазирования двух сигналов

Значения весовых коэффициентов $W1$ и $W2$, обеспечивающих синфазность сигналов $X(t)$ и $Z(t)$, можно найти из известного в теории адаптивной обработки сигналов матричного уравнения Винера–Хопфа [13], которое для рассматриваемой процедуры синфазирования имеет вид

$$\begin{pmatrix} W1 \\ W2 \end{pmatrix} = \begin{pmatrix} \langle Z1(t) \cdot Z1(t) \rangle & \langle Z1(t) \cdot Z2(t) \rangle \\ \langle Z2(t) \cdot Z1(t) \rangle & \langle Z2(t) \cdot Z2(t) \rangle \end{pmatrix}^{-1} \begin{pmatrix} \langle X(t) \cdot Z1(t) \rangle \\ \langle X(t) \cdot Z2(t) \rangle \end{pmatrix}. \quad (16)$$

Тогда сигнал $ZC(t)$, синфазный с сигналом $X(t)$, определяется соответственно:

$$ZC(t) = W1 \cdot Z1(t) + W2 \cdot Z2(t). \quad (17)$$

При цифровой реализации процедуры синфазирования весовые коэффициенты $W1$ и $W2$, полученные на основе уравнения (16), определяются зависимостями:

$$W1 = \frac{\sum_{n=1}^L Z1(n)X(n)}{\sum_{n=1}^L Z1(n)Z1(n)}; \quad W2 = \frac{\sum_{n=1}^L Z2(n)X(n)}{\sum_{n=1}^L Z2(n)Z2(n)}, \quad (18)$$

где $X(n)$, $Z1(n)$, $Z2(n)$ – отсчёты сигналов $X(t)$, $Z1(t)$, $Z2(t)$ после синхронной дискретизации; L – количество отсчётов в выборке усреднения.

Форма записи цифрового синфазированного сигнала $ZC(n)$ аналогична соотношению (17), но применительно к отсчётам квадратурных составляющих вспомогательного сигнала:

$$ZC(n) = W1 \cdot Z1(n) + W2 \cdot Z2(n). \quad (19)$$

4. Модифицированный метод измерения разности фаз сигналов

С учётом представленных в разделах 2, 3 сведений разработанный квадратурный метод также удобно представить в виде определённой последовательности действий и операций.

Сигналы $X(t)$, $Y(t)$ выбранной низкой промежуточной частоты $f_{\text{пр}}$ (с априорно неизвестным отношением сигнал/квазибелый шум) и квадратурные составляющие $Z1(t)$, $Z2(t)$ вспомогательного сигнала промежуточной частоты (с большим отношением сигнал/квазибелый шум) подвергаются синхронной дискретизации с кратным четырём количеством отсчётов k на периоде колебаний и общим числом отсчётов L в выборке, формируются четыре массива выборок $X(n)$, $Y(n)$, $Z1(n)$, $Z2(n)$:

$$X(t) \rightarrow X(n); Y(t) \rightarrow Y(n); Z1(t) \rightarrow Z1(n); Z2(t) \rightarrow Z2(n); n \in [1 \div L], \quad (20)$$

где n – порядковый номер отсчёта в выборке, при этом параметры k и L выбираются равными $k = 3.6 \cdot 10^3$, $L = 100k$.

Далее выполняются следующие программные операции.

Вычисляются весовые коэффициенты $W1$, $W2$, $W3$, $W4$, необходимые для синфазирования двух пар сигналов $X(t)$ и $Z(t)$, $Y(t)$ и $Z(t)$:

$$W1 = \frac{\sum_{n=1}^L Z1(n)X(n)}{\sum_{n=1}^L Z1(n)Z1(n)}; W2 = \frac{\sum_{n=1}^L Z2(n)X(n)}{\sum_{n=1}^L Z2(n)Z2(n)}; W3 = \frac{\sum_{n=1}^L Z1(n)Y(n)}{\sum_{n=1}^L Z1(n)Z1(n)}; W4 = \frac{\sum_{n=1}^L Z2(n)Y(n)}{\sum_{n=1}^L Z2(n)Z2(n)}. \quad (21)$$

Рассчитываются и формируются два массива $RX(n)$, $RY(n)$ отсчётов сигналов, синфазных с сигналами $X(n)$, $Y(n)$ соответственно, имеющих в результате синфазирования большие отношения сигнал/шум:

$$RX(n) = W1 \cdot Z1(n) + W2 \cdot Z2(n); RY(n) = W3 \cdot Z1(n) + W4 \cdot Z2(n); n \in [1 \div 1.25 \cdot k]. \quad (22)$$

Осуществляется нормирование отсчётов массивов $RX(n)$, $RY(n)$ путём их представления своими знаками:

$$RX(n) = \begin{cases} 1, & \text{если } RX(n) > 0 \\ -1, & \text{если } RX(n) < 0 \end{cases}; RY(n) = \begin{cases} 1, & \text{если } RY(n) > 0 \\ -1, & \text{если } RY(n) < 0 \end{cases}; n \in [1 \div 1.25 \cdot k]. \quad (23)$$

Вычисляются корреляционные коэффициенты β и γ :

$$\beta = \frac{1}{k} \sum_{g=1}^k RX(g)RY(g); \gamma = \frac{1}{k} \sum_{p=1+0.25k}^{1.25k} RX(p)RY(p-0.25k). \quad (24)$$

Рассчитывается искомая разность фаз $\Delta\varphi_{\text{и}}$ с использованием формулы:

$$\Delta\varphi_{\text{и}} = \begin{cases} 90(1-\beta), & \text{если } \gamma \geq 0 \\ -90(1-\beta), & \text{если } \gamma < 0 \end{cases} [\text{град.}]; \Delta\varphi_{\text{и}} \in [-180^\circ \div +180^\circ]. \quad (25)$$

5. Эффективность процедуры синфазирования сигналов

Повышение эффективности цифрового квадратурного метода измерения разности фаз при введении процедуры синфазирования оценивалось путём сравнения максимальных абсолютных погрешностей измерений обычного квадратурного метода и его модификации для следующих, интересных с точки зрения практического использования, параметров:

– изменения отношения сигнал/квазибелый шум трактов преобразования частоты в диапазоне $N \in [-20 \div +10]$ дБ;

– трёх значений коэффициента широкополосности ($N = 5, 15, 50$) трактов преобразования частоты;

– значений измеряемой разности фаз, при которой абсолютная погрешность измерения на реализации квазибелого шума максимальна, из интервала $\Delta\varphi_{\text{и}} \in [-165^\circ \div +165^\circ]$;

– осреднения измеренных разностей фаз $\Delta\varphi_{\text{и}}$ по 25 независимым реализациям нормального квазибелого шума с нулевым средним значением и заданной дисперсией.

На рис. 5 представлены зависимости модуля максимальной абсолютной погрешности измерения разности фаз сравниваемых методов от отношения сигнал/квазибелый шум трактов преобразования частоты, полученные в результате моделирования.

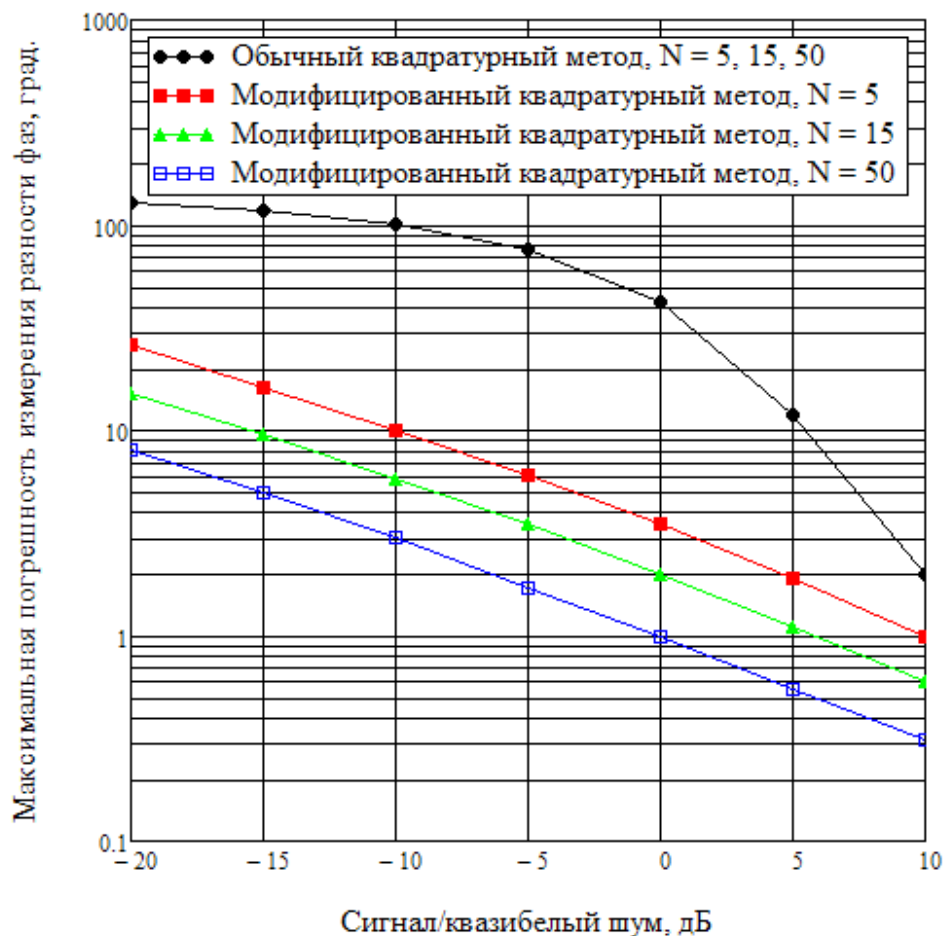


Рис. 5. Сравнительная эффективность квадратурных методов

Математическое моделирование сравниваемых методов позволило сделать следующие полезные в плане технической реализации выводы:

- погрешности измерений разности фаз модифицированного квадратурного метода в диапазоне отношений сигнал/квазибелый шум трактов преобразования частоты $N \in [-20 \div +20]$ дБ всегда меньше погрешностей обычного квадратурного метода;
- погрешности измерений разности фаз обычного квадратурного метода не зависят от величины полосы пропускания трактов преобразования частоты и определяются только отношением сигнал/квазибелый шум в указанных трактах;
- погрешности измерений разности фаз модифицированного квадратурного метода уменьшаются с увеличением коэффициента широкополосности N (полосы пропускания при фиксированной промежуточной частоте) трактов преобразования частоты;
- точности измерений разности фаз в модифицированном квадратурном методе по сравнению с обычным квадратурным методом для параметров $N = 50$, $N \in [-10 \div +10]$ дБ увеличиваются от 7 до 30 раз в зависимости от отношения сигнал/квазибелый шум трактов преобразования частоты.

6. Заключение

Разработанный цифровой модифицированный квадратурный метод измерения разности фаз гармонических сигналов позволяет повысить точность измерений в линейных трактах с малыми отношениями сигнал/квазибелый шум за счёт выбора необходимого соотношения полосы пропускания/промежуточная частота трактов преобразования частоты и перехода от измерения разности фаз сигналов с малыми отношениями сигнал/шум к измерению разности фаз сигналов с заведомо большими отношениями сигнал/шум, обеспечиваемому процедурой синфазирования вспомогательного сигнала и преобразованных измерительных сигналов исследуемых трактов.

Ожидаемая реализация метода ориентирована на современную элементную базу и существующую востребованность в новейших разработках гидроакустических информационно-навигационных систем.

Литература

1. Чмых М. К. Цифровая фазометрия. М.: Радио и связь, 1993. 184 с.
2. Webster J. G. Electrical measurement, signal processing, and displays. Boca Raton: CRC Press, 2004. 723 p.
3. Метрология и радиоизмерения / под ред. В. И. Нефедова. М.: Высшая школа, 2006. 526 с.
4. Mahmud S. M. Error analysis of digital phase measurement of distorted waves // IEEE Trans. on Instrumentation and Measurement. 1989. V. 38, № 1. P. 6–9.
5. Liang Y. R., Duan H. Z., Yeh H. C., Luo J. Fundamental limits on the digital phase measurement method based on cross-correlation analysis // Rev. Sci. Instrum. 2012. V. 83, № 9. P. 95–110.
6. Mahmud S. M. High precision phase measurement using reduced sine and cosine tables // IEEE Trans. on Instrumentation and Measurement. 1990. V. 39, № 1. P. 56–60.
7. Mahmud S. M. High precision phase measurement using adaptive sampling // IEEE Trans. on Instrumentation and Measurement. 1989. V. 38, № 5. P. 954–960.
8. Гоноровский И. С. Радиотехнические цепи и сигналы. М.: Дрофа, 2006. 722 с.
9. Sedlacek M., Krumpholtz M. Digital measurement of phase difference a comparative study DSP algorithms // Metrology and Measurement Systems. 2005. V. XII, № 4. P. 427–449.
10. Дамдинова Д. Б., Полетаев А. С., Ченский А. Г. Сравнение точности методов вычисления разности фаз квазигармонических сигналов // Вестник СибГУТИ. 2016. № 2. С. 87–97.

11. *Шиховцев И. В., Якубов В. П.* Статистическая радиофизика: курс лекций. Новосибирский государственный университет, 2011. 157 с.
12. *Тихонов В. И., Харисов В. Н.* Статистический анализ и синтез радиотехнических устройств и систем. М.: Радио и связь, 1991. 608 с.
13. *Уидроу Б., Стирнз С.* Адаптивная обработка сигналов / пер. с англ. под редакцией В. В. Шахгильдяна. М.: Радио и связь, 1989. 440 с.

Статья поступила в редакцию 20.03.2019.

Арсентьев Виктор Георгиевич

к.т.н., ведущий научный сотрудник научно-технического центра специализированных информационных систем СибГУТИ (630008, Новосибирск, ул. Бориса Богаткова, 51), тел. (383) 2-693-938, e-mail: viktor.arsentev.51@mail.ru.

Криволапов Геннадий Илларионович

к.т.н., доцент, заведующий лабораторией, руководитель научно-технического центра специализированных информационных систем СибГУТИ, тел. (383) 2-693-942, e-mail: krivolapov@sibsutis.ru.

**Measurement of phase difference of harmonic signals in linear paths
with quasi-white noise**

V. Arsent'ev, G. Krivolapov

A quadrature method for measuring the phase difference of harmonic signals using frequency conversion is considered. It is proposed to introduce into the algorithmic structure of the method the common mode procedure of signal. The efficiency of the developed modified quadrature method is estimated.

Keywords: the phase difference, quadrature method, common mode signals.

Ретроанализ менеджмента эксплуатации систем электросвязи

А. И. Ициксон

Исследование посвящено определению выбора стратегий эксплуатации систем и средств электросвязи с происшедшими изменениями технологии изготовления оборудования. Существующие исследования не увязаны со старением изделий в зависимости от технологий изготовления радиоэлектронной аппаратуры. Предложена модель корреляции стратегии эксплуатации электросвязи с совершенствованием технологии. На базе методики Госстандарта выделены основные стратегии эксплуатации; на математико-экономических моделях стратегий выполнено сопоставление по критерию минимума расходов/максимума доходов. Определено изменение скорости старения. Проведен эволюционный анализ совершенствования технологии с 1960-х гг. Проведены натурные исследования скорости дрейфа (старения) параметров, определяющих работоспособность устройств и систем электросвязи для разных элементных компонентов: на базе ламповой техники, полупроводников, интегральных микросхем, работавших в 1970–1990-х годах. Выполнено сравнение стратегий эксплуатации на нынешнем уровне развития техники. Доказано, что дальнейшего изменения стратегий эксплуатации не будет. Дальнейшее совершенствование технологии изготовления будет сопровождаться количественными изменениями.

Ключевые слова: электросвязь, менеджмент эксплуатации, стратегии эксплуатации, скорости изменения параметров, критерий минимума расходов на эксплуатацию.

1. Введение

Технология эксплуатации оборудования электросвязи меняется в течение последних 60 лет. Ее изменения вызваны появлением техники принципиально нового назначения с изменением элементной базы, а также повышением надежности изделий на меняющейся элементной базе. Эффективность применения систем по назначению определяется уровнем надежности и применяемой стратегией эксплуатации (СЭ). Актуальность и значимость проблемы во всех сферах деятельности определили исследования и публикации, начиная с теоретического описания возможных вариантов математических методов в теории надежности более полувека назад в середине 1960-х [1, 2, 3] и решения широкого диапазона прикладных задач надежности для эксплуатации автоматики и автоматизированных систем управления [4], для судовой радиоэлектронной аппаратуры [5], в военном деле [6] и для ракетной техники [7]. Теоретически авторы рассматривают различные физические модели старения [8–12]. Общий подход учета старения технических изделий учтен в методике [13]. Большинство исследований являются специфичными, носят сугубо теоретический характер исходя из частных случаев оптимизации. Другой класс исследований увязывает вопросы надежности и экономики, прогнозирования изменений параметров для разной элементной базы оборудования и его эксплуатации в конкретных условиях эксплуатации изделий различного назначения [14–23]. Цель данной работы – определить тенденции изменения стратегии эксплуатации в соответствии с изменениями технологии изготовления и старения средств электросвязи.

2. Основные этапы развития технологии изготовления

В СССР радиолампы массово использовались в XX веке как активные элементы электронной аппаратуры для электросвязи (усилители, генераторы, детекторы, переключатели). «Золотая эра» ламповой схемотехники пришлась на 1935–1950 годы. Министерство промышленности средств связи выпускало устройства на радиолампах и в 1960-е годы. Первая отечественная микросхема была создана в 1961 году, серийное производство гибридных толстоплёночных интегральных микросхем – с 1965 года, заводское производство интегральных микросхем – с 1967 г. Большой ассортимент аналоговых интегральных микросхем был получен к концу 1970-х годов. Их применение позволило увеличить надёжность устройств, упростить наладку оборудования, часто даже исключить необходимость технического обслуживания в процессе эксплуатации. В 1970–1980-х годах минимальный контролируемый размер серийно производимых микросхем был уменьшен в четыре раза.

Развитие систем передачи данных на большие расстояния упиралось в фундаментальные ограничения: для электрических систем это явление затухания сигнала на определённом расстоянии, для сверхвысокочастотных систем – несущая частота. Во второй половине XX века решение было найдено – оказалось, что передача сигнала с помощью света гораздо эффективнее как электрического, так и СВЧ-сигнала. Существовало две глобальных проблемы при разработке оптических систем передачи данных: источник света и носитель сигнала. Первая решилась с изобретением лазеров в 1960 году, вторая – с появлением высококачественных оптических кабелей в 1970 году. Первый в мире практически применимый светодиод разработан в 1962 году, полупроводниковый диод Ганна для генерации и преобразования колебаний в диапазоне СВЧ – в 1963 г., первый в мире жёлтый светодиод с улучшением яркости в 10 раз – в 1972 году. Светодиоды оставались чрезвычайно дорогими, их практическое применение было ограничено до 1971 г., когда было организовано массовое производство светодиодов, применимых в индикаторах. В 1976 году создан первый в мире высокоэффективный светодиод высокой яркости для передачи данных по волоконно-оптическим линиям связи (ВОЛС). В 1980-х изобретен дешёвый синий светодиод, к 1993 г. компании «Nichia» удалось начать промышленный выпуск синих светодиодов нового типа. После интенсивных исследований в период с 1975 по 1980 год появилась первая коммерческая ВОЛС, использовавшая полупроводниковый лазер на основе арсенида галлия. Скорость передачи систем первого поколения составлял 45 Мбит/с, расстояние между повторителями – 10 км. Второе поколение волоконно-оптических систем было разработано для коммерческого использования в начале 1980-х гг.; уже в 1987 году эти системы работали на скорости до 1.7 Гбит/с при расстоянии между повторителями 50 км. Прокладка первой в мире трансокеанской волоконно-оптической линии связи была завершена в 1988 году (между Японией и США), её длина составила около 10 тысяч километров; к 2003 году при применении технологии спектрального уплотнения была достигнута скорость передачи 10.92 Тбит/с. В 2013 году ученые из компании «Bell» протестировали технологию шумоподавления с передачей 400 Гбит/с по оптоволокну на 12 800 км без повторителей сигнала. В среднем внедрение лидерами индустрии новых технологий происходило примерно каждые два года.

В данной статье рассмотрены этапы с условными временными границами: первый этап: 1960–1970-е гг. – электронные лампы; второй этап: 1970–1980-е гг. – дискретные транзисторы и полупроводниковые диоды; третий этап: 1980–1990-е гг. – интегральные схемы (микросхемы) и волоконно-оптические линии связи. К 1980-м годам интегральные микросхемы благодаря своей миниатюрности, экономичности, устойчивости к механическим воздействиям и невысокой стоимости практически полностью вытеснили радиолампы из малосигнальной электроники; ВОЛС массово вытеснили медножильные линии с магистральных направлений, затем из распределительных сетей и абонентских линий связи. Последующие годы характеризовались увеличением ресурса изделий, но качественных скачков технологии изготовления или появления новых принципов, новых физических явлений для работы изделий электро-

связи не было. Краткий экскурс в историю совершенствования технологии производства показывает, что электросвязь, компьютерная техника и радиоэлектроника развивались в основном в последние 50 лет, многие виды систем связи и оборудование массового потребления появились в последние десятилетия.

3. Методология

Уровень эффективности использования изделий по назначению в значительной степени определяется уровнем надежности (поддержания заданных параметров функционирования во времени) [24] и выбранной стратегией эксплуатации. Стратегия эксплуатации – это система правил, по которым осуществляется управление их техническим состоянием, принимается решение о моменте начала и объеме операций обслуживания и ремонта всех видов, в том числе затратами на техническое обслуживание и ремонт (ТО и Р) [25]. За прошедшие годы интенсивно велись исследования по проблеме надежности, выделился ряд научных направлений. Появилось значительное количество монографий по надежности, авторы которых основное внимание уделяли углубленному рассмотрению сравнительно узкого круга задач, тогда как проблема надежности является комплексной, системотехнической. Рост значения проблемы надежности связан с особенностями развития современной техники. Во-первых, существует стремление к подробному планированию хода производственных процессов, которые становятся все более сложными. Во-вторых, все больше распространяется автоматизация различных процессов. В-третьих, автоматизированные системы выполняют все более ответственные задачи. Кроме того, современные конструкции создаются в очень короткий срок и быстро морально стареют, что затрудняет накопление навыков их проектирования, изготовления и эксплуатации. Обычно более заметны непосредственные результаты недостаточной надежности, связанные с полным или частичным невыполнением требуемых от технического объекта функций. Все это обуславливает необходимость общего подхода к снижению расходов и повышению надежности всех систем независимо от их устройства и назначения. Установилась традиция проводить инженерное исследование надежности технических объектов по данным об отказах. Постепенные отказы (параметрические, плавные) связаны с износом деталей, старением материалов и разрегулированием устройств. Параметры объекта могут достигнуть критических значений, при которых его состояние считается неработоспособным, т.е. происходит отказ. В настоящее время нет однозначного физического толкования происхождения применяемых распределений наработки до отказа. Вместе с тем во многих случаях за время эксплуатации или испытаний на надежность успевает отказать лишь незначительная часть первоначально имевшихся объектов. Поэтому значения числовых характеристик, найденные в результате обработки опытных данных, сильно зависят от типа предполагаемого распределения наработки до отказа. Выбранному теоретическому распределению наработки до отказа должна соответствовать определенная модель приближения объекта к отказу. Желательно, чтобы эта модель обобщала основные особенности физических процессов приближения к отказам. Одной из основных трудностей при исследовании вопросов надежности является ограниченность статистической информации о процессе эксплуатации объектов. Эти сведения могут быть получены путем наблюдения за работой объектов либо в условиях реальной эксплуатации, либо при испытаниях. Данные испытаний не могут полностью заменить эксплуатационные данные, а реальная эксплуатация представляет собой недостижимый по своим масштабам в лабораторных условиях эксперимент. До проведения испытаний проводятся аппроксимация имеющихся экспериментальных данных каким-либо теоретическим распределением и проверка статистической гипотезы о том, что принятое теоретическое распределение не противоречит экспериментальному распределению. При различных предположениях о виде распределения наработки до отказа значения средней наработки до отказа объектов, вычисленные по одним и тем же опытным данным, могут отличаться в сотни раз. Изложенные причины определяют целесообразность применения во многих случаях для описания процессов изна-

шивания линейных случайных процессов, все реализации которых являются прямыми линиями. При линеаризации процесса изнашивания каждая реализация этого процесса заменяется секущей прямой, т.е. рассматриваются лишь главные, линейные части монотонных случайных процессов изменения параметра. Линейные случайные процессы являются очень удобной моделью процессов изнашивания во многих задачах. Эти процессы близки к встречающимся в действительности процессам, очень просто описывают основные особенности процессов изменения параметра, требуют минимального количества экспериментальных данных для вычисления характеристик случайного процесса и дают возможность наиболее просто исследовать надежность элементов при изменении их свойств. На основании представленных критериев применительно к законам изменения параметров и показателей затрат на эксплуатацию возможно выполнить технико-экономическое сравнение стратегий. Экономическая эффективность системы зависит от соотношения прибыли и расходов на ее эксплуатацию. С прикладной точки зрения данный факт влияет на выбор оптимальной стратегии эксплуатации.

Систематизации и оптимизации моделей СЭ более полвека продолжают посвящать множество работ. Разработки стратегий восстанавливаемых изделий методически обобщены ВНИИ стандартизации описанием математико-экономических моделей в [19]. На основании представленных критериев применительно к законам изменения параметров и показателей затрат на ТО и Р возможно выполнить экономическое сравнение стратегий. Методика распространяется на «...устройства, эффект от эксплуатации которых пропорционален наработке, а обслуживание сводится к предупредительным заменам, заменам вследствие отказа и периодическим проверкам работоспособности с последующим проведением технического обслуживания, ремонта и диагностирования». Выбранный в методике критерий оптимизации – максимум эффекта на единицу затрат, необходимых для его достижения, является общим для следующих критериев оптимизации: минимум суммарных затрат и удельных затрат на эксплуатацию устройства, максимум коэффициента технического использования, максимум коэффициента готовности. В методике формализованы шесть стратегий СЭ.

Критерии оптимизации K_i стратегий эксплуатации основываются на базовых технико-экономических показателях с учетом: средних затрат на единицу времени при проведении работ по обслуживанию (предупредительной замены β_1 , технической диагностики β_2 , замены вследствие отказа β_3); средней продолжительности проведения этих работ (предупредительной замены t_1 , технической диагностики t_2 и замены вследствие отказа t_3); среднего ущерба ε_1 за единицу времени простоя (или невыполнения устройством единицы работы) и среднего ущерба ε_2 от отказа устройства; периодичности проведения предупредительных замен T_1 и проверок работоспособности устройства T_2 ; функции восстановления Ω , описываемой интегральным уравнением Вольтерра второго рода с разностным ядром [1]; t_x – средняя наработка на отказ; $\omega_{вн}$ – параметр потока внезапных отказов; ω_n – параметр потока постепенных отказов; β – относительный ресурс; α – относительная скорость изменения параметров, которая обратно пропорциональна коэффициенту вариации скорости изменения параметров (коэффициент вариации – это отношение среднеквадратического отклонения к среднему значению), $P(t)$ – вероятность безотказной работы за время t . Сравнение СЭ выполнено в [20].

1. Стратегия эксплуатации СЭ1 по мере возникновения отказов.

Критерий $K_1 = (\beta_3 + \varepsilon_1 \cdot t_3 + \varepsilon_2) / t$.

Описание. Замены вследствие отказа, который обнаруживается мгновенно. Обслуживание при отказе – это когда ремонтные работы будут выполняться только в случае достижения критического состояния оборудования, при котором оно уже не может выполнять заданные функции, то есть теряет работоспособность. Стратегия применима для устройств, где можно допускать потерю работоспособности, аварийные ситуации, серьезные поломки с длительным устранением, в результате – излишние расходы на ликвидацию последствий и потери от остановки производства. Требуется формирование завышенного резерва материальных ресурсов.

2. Стратегия предупредительных замен СЭ2 вследствие отказа без перенесения времени проведения очередной предупредительной замены. Иные наименования стратегии: по ресурсу, планово-предупредительных ремонтов (ППР).

$$\text{Критерий } K_2 = [\beta_1 + \varepsilon_1 \cdot t_1 + (\beta_3 + \varepsilon_1 \cdot t_3 + \varepsilon_2) \cdot \varrho] / T_1.$$

Описание. Техническое обслуживание, предусмотренное в нормативно-технической или эксплуатационной документации. Выполняется с периодичностью и в объеме, независимо от технического состояния изделия в момент начала технического обслуживания. Это технология предупредительного технического обслуживания и ремонтов исходя из статистических сведений о сроках службы оборудования. Наибольшее распространение стратегия ППР получила при плановой экономике. Стратегией ППР и сегодня пользуются на многих предприятиях, прежде всего для обслуживания ключевых стратегических объектов, остановка которых может причинить вред окружающей среде, жизни и здоровью людей или стоимость отказа очень велика. Благодаря систематическому проведению техобслуживания и ремонта и своевременной замене достигается более качественная и длительная эксплуатация оборудования. Недостаток ППР – нередко ремонтируются фактически исправные объекты, а также производится обязательная замена деталей независимо от их оставшегося ресурса. В результате этого эксплуатационные затраты становятся неоправданно высоки. Сопоставление стратегий позволяет определить предельное значение периода обслуживания, когда оно экономически не выгодно, но необходимо по условиям безопасности: $K_2 = K_1$ или $T_1 = (C_{TO} + \varrho) \cdot t$, где $C_{TO} = (\beta_1 + \varepsilon_1 \cdot t_1) / (\beta_3 + \varepsilon_1 \cdot t_3 + \varepsilon_2)$ – коэффициент целесообразности ППР. Граничные условия при $\alpha \rightarrow 0$ зависят от параметра потока внезапных отказов. Для экспериментально определенного диапазона $\beta \cdot \omega_{вн} \in (0...2.5)$ при $C_{TO} \geq (0.1 \dots 0.3)$ эффективна СЭ1. Соответственно, чем меньше C_{TO} , тем выгоднее стратегия СЭ2. То есть при незначительных затратах C_{TO} , при уменьшении α , при уменьшении параметра потока внезапных отказов $\omega_{вн}$ экономически эффективно ТО с периодичность T_1 . При больших C_{TO} , α , $\omega_{вн}$ более эффективна СЭ1 по мере возникновения отказов. При поступлении изделий в эксплуатацию существует период «приработки», в течение которого параметр потока внезапных отказов может превышать значение в период «нормальной эксплуатации» в несколько раз. Поэтому часто эффективной является СЭ по мере возникновения отказов с последующим переходом на СЭ предупредительных замен. Для выявленного диапазона $\alpha \in (0.01...1.0)$ автором описан запатентованный алгоритм нахождения периода T_1 [11].

3. Стратегия СЭЗ – частный случай стратегии СЭ2, когда выполняется перенос времени очередной предупредительной замены.

$$\text{Критерий } K_3 = (\beta_1 + \varepsilon_1 \cdot t_1) \cdot P(T_1) + (\beta_3 + \varepsilon_1 \cdot t_3 + \varepsilon_2) \cdot [1 - P(T_1)] \cdot \int_0^{T_1} P(t) dt.$$

Описание. Стратегия характеризуется оцениванием вероятности отказов. Степень «важности» оборудования и вероятность его отказа определяют уровень риска при эксплуатации оборудования. Исходя из уровня риска, назначаются сроки и объёмы ТО и Р, расставляются приоритеты выполнения ремонтных работ. В результате повышается надёжность и безопасность, понижаются сопутствующие затраты. Стратегия направлена на минимизацию потерь от: непродуктивной работы оборудования, простоев, ненужных перемещений обслуживающего персонала, недостатков обеспечения ресурсами, повторных переделок ремонтных работ, нерационального использования ресурсов, неэффективного управления информацией. Для данной стратегии определяются наиболее эффективные мероприятия, обеспечивающие безопасность и надёжность оборудования (снижение рисков) с минимизацией сопутствующих затрат. Требование $K_2 > K_3$ практически всегда выполняется, то есть перенос времени замены всегда экономически эффективен.

4. Стратегия СЭ4 проверки и замены вследствие отказа при проверке работоспособности устройства. Иное наименование: с техническим диагностированием (ТД).

$$\text{Критерий } K_4 = [\beta_3 + \varepsilon_1 \cdot t_3 + \varepsilon_2 + (\beta_2 + \varepsilon_1 \cdot t_2 + \varepsilon_1 \cdot T_2) \cdot \sum_{n=0}^{\infty} P(n \cdot T_2)] / (t_x - \varepsilon_1).$$

Описание. Периодическая проверка работоспособности устройства. При обнаружении отказа – замена устройства. Контроль технического состояния выполняется с установленными периодичностью и объемом, а объем остальных операций определяется техническим состоянием изделия. Техническое диагностирование всегда экономически эффективно, позволяя уточнять объем ТО и Р, но, исходя из минимума затрат, это должна быть встроенная система функционального диагностирования. При выполнении данного требования именно данная стратегия будет оптимальна для устройств и систем, обеспечивающих безопасность. При отсутствии встроенной, сравнительно дешевой, диагностики оптимальна стратегия эксплуатации по ресурсу СЭ2: $K_2 \geq K_4$ при $T_1 \gg T_2$: $\sum_{n=0}^{\infty} P(n \cdot T_2)/t_x \approx 1/T_2$, тогда $(C_{то} + \Omega) / T_1 - 1/t_x \geq C_{тд} / T_2$, где $C_{тд} = (\beta_2 + \varepsilon_1 \cdot t_2) / (\beta_3 + \varepsilon_1 \cdot t_3 + \varepsilon_2)$ – коэффициент целесообразности ТД. Данное выражение определяет величину допустимых удельных затрат на систему технического диагностирования в зависимости от показателей надежности оборудования и уровня затрат на его плановое обслуживание. Тогда для оптимального значения периодичности ТД имеем $T_2 = C_{тд} / (\omega_n + \omega_{вн} - 1/t_x)$.

5. Прогнозная, или проактивная, стратегия СЭ5 как усовершенствованный вариант СЭ6. Критерий

$$K_5 = [(\beta_1 + \varepsilon_1 \cdot t_1) \cdot P(T_1) + (\beta_3 + \varepsilon_1 \cdot t_3 + \varepsilon_2) \cdot [1 - P(T_1)] + (\beta_2 + \varepsilon_1 \cdot t_2 + \varepsilon_1 \cdot T_2) \sum_{n=0}^{\infty} P(n \cdot T_2)] / \int_0^{T_1} P(t) dt - \varepsilon_1.$$

Описание. Проверки и предупредительные замены вследствие отказа. Периодическая проверка работоспособности устройства; вместе с плановой проверкой выполнять предупредительную замену, если до этого не был обнаружен отказ. При обнаружении отказа заменить устройство с перенесением времени очередной предупредительной замены. Отсчет количества проверок производится от последней предупредительной замены. Это более продвинутый подход, совмещающий преимущества планирования по текущему состоянию с преимуществами ППР. Суть проактивной стратегии – снижение скорости развития или устранение неисправности, выявленной при мониторинге технического состояния оборудования. В проактивной стратегии ключевым моментом является диагностика состояния оборудования, которая может выполняться путем: визуального осмотра, мониторинга технических параметров, акустической и вибрационной диагностики. Решение о ремонте принимается, если неудовлетворительное состояние одного элемента оборудования начинает негативно влиять на состояние других элементов. Применение проактивной стратегии увеличивает срок службы, исключает вторичные повреждения в связи с первичной поломкой путем немедленной реакции на первичную поломку, сокращает общие затраты на эксплуатацию, снижает риск аварийного отказа оборудования, повышает коэффициент исправности оборудования.

6. Стратегия по техническому состоянию СЭ6.

$$\text{Критерий } K_6 = [(\beta_1 + \varepsilon_1 \cdot t_1) \cdot P(T_1) + (\beta_3 + \varepsilon_1 \cdot t_1 + \varepsilon_2) \cdot [1 - P(T_1)] + \varepsilon_1 \cdot T_1] / \int_0^{T_1} P(t) dt - \varepsilon_1.$$

Описание. Предупредительные замены и замены вследствие отказа, замены периодически по календарному принципу. Техническое обслуживание, предусмотренное в нормативно-технической или эксплуатационной документации и выполняемое по результатам непрерывного контроля технического состояния изделия с применением диагностической аппаратуры для мониторинга оборудования в режиме реального времени. Область применения стратегии эксплуатации целесообразно отнести к объектам: отказы которых не влияют непосредственно на безопасность; надежность которых позволяет обеспечить выполнение требований по экономической эффективности; обладающим высокой эксплуатационной технологичностью; затраты на эксплуатацию которых до отказа при обслуживании с контролем уровня надежности не превышают затрат на ППР; имеющим индикацию отказов средствами контроля с минимальными трудовыми затратами в заданное время. При использовании этой стратегии благо-

даря постоянному мониторингу риск аварийного отказа или серьезного ухудшения работоспособности оборудования минимизируется. Такой подход уменьшает расходы, сводит к минимуму число непредвиденных отказов. СЭ6 по техническому состоянию эффективнее СЭ2 по ресурсу при выполнении неравенства $f(T_1) - (1/t_x - \omega_{вн}) \geq C_{тд} / T_2$.

Стратегия СЭ6 для объектов, отказы которых не влияют непосредственно на безопасность, всегда экономически уступает стратегии СЭ2, когда вместо календарного принципа планирования учитывается фактический ресурс объекта. Стратегия СЭ3 всегда эффективнее стратегии СЭ2, поскольку происходит экономия в связи с переносом времени очередного планового ТО после возникновения отказа. Таким образом, из трех стратегий СЭ6, СЭ2 и СЭ3, связанных с плановым ТО, всегда экономически эффективна стратегия СЭ3.

Возможность сопоставления СЭ согласно описанным критериям появляется для конкретной модели, описывающей изменение параметров. Эксплуатация системы описывается функцией восстановления $\Omega(t)$ или ее дифференциальной характеристикой – параметром потока отказов $\omega(t)$. Эти функции удовлетворяют интегральному уравнению Вольтерра второго порядка с разностным ядром. Для его решения необходимо конкретизировать, каким распределением описывать функцию восстановления. Расчет показателей надежности по постепенным отказам производится по данным дрейфа определяющих параметров изделия, что позволяет интегрально учитывать изменение физико-химических свойств элементарных компонентов и, соответственно, влияние функциональной структуры исследуемого изделия на показатели надежности с точки зрения физического старения элементов (постепенных отказов) при описании изменения в рамках допусков изменения параметров, определяющих работоспособность изделия.

В 1960–1970-е годы на практике для радиоэлектронной и электромеханической аппаратуры рассматривалось распределение с $\alpha > 2.5$. Теоретически для диапазона значений $\alpha \in (2...15)$ было предложено приближенное решение интегрального уравнения Вольтерра второго рода с разностным ядром [11, 12]. При этом выдвинута гипотеза распределения скорости изменения параметра по нормальному закону. Нормальный закон (закон Гаусса) широко используется при оценке надежности изделий, на надежность которых воздействует ряд случайных факторов, каждый из которых незначительно влияет на результирующий эффект (нет доминирующих факторов). В этом случае распределение скорости изменения параметров подчинено нормальному закону, а плотность характеризуется альфа-распределением. При описании параметра потока отказов альфа-распределением статистически показано, что при совершенствовании технологии производства происходит уменьшение показателя α , наиболее полно описывающего изменение параметров в условиях эксплуатации.

4. Натурные исследования старения средств и систем электросвязи

Предложена и подтверждена статистически следующая гипотеза (подтвержденная на изделиях машиностроения [11] и автоматики [12]): средняя скорость изменения параметров, определяющих работоспособность изделий, является квазипостоянной. Распределение скоростей однотипных изделий подчинено нормальному закону. Функция распределения тогда описывается альфа-распределением. Распределение характеризуется показателем α – относительной скоростью изменения параметров, который обратно пропорционален коэффициенту вариации скорости изменения параметров, и показателем β – относительным запасом долговечности – ресурсом, равным отношению значения среднего ресурса к коэффициенту вариации скорости дрейфа (изменения) параметров.

Исторически менялась элементная база, на которой построены системы электросвязи. Анализ процессов старения показал, что переход от ламповых элементов к полупроводникам, затем к микросборкам и дальнейшее совершенствование микросхем, так же как и изменение технологии производства волоконно-оптических кабелей сопровождалось снижением средней величины при увеличении дисперсии скорости старения [26]. Проверка гипотезы нормального

распределения скорости дрейфа параметров, определение средней скорости и оценка средне-квадратического отклонения осуществлялись в результате обработки данных для отдельных параметров по стандартной методике, значения полей допуска – допустимых изменений параметров взяты из стандарта [27] и технических условий на соответствующие изделия. В качестве прогнозируемых выбирались параметры, проверка и регулировка которых выполняется в соответствии с инструкциями по эксплуатации заводов-изготовителей. Методология расчета основывается на формализованном подходе [13]. Репрезентативность выборки определялась в соответствии с [28]. В результате многолетних исследований (натурного эксперимента) на репрезентативных выборках в условиях эксплуатации подтверждено, что скорость дрейфа параметров аппаратуры электросвязи квазипостоянна, ее распределение подчинено нормальному закону с большими коэффициентами вариации (малыми значениями α) [20, 26]. Для изделий электросвязи на радиолампах получено $\alpha = 0.7$ (разброс по отдельным параметрам, определяющим работоспособность изделия в целом, – от 0.3 до 1.4); переход на полупроводниковые микросхемы сопровождался снижением $\alpha \in (0.25 \dots 0.40)$ (разброс для определяющих параметров – от 0.2 до 1.0), схемная модификация оборудования – переход на аналоговые интегральные микросхемы определила дальнейшее снижение до $\alpha \in (0.01 \dots 0.06)$. Экспериментально подтверждено [26], что каждое качественное изменение технологии сопровождалось и увеличением (более трехкратного) нормированной наработки на отказ β (максимум функции восстановления уменьшается вдвое при уменьшении α от 1 до 0; время наступления максимума увеличивается с $0.5/\beta$ до $0.707/\beta$). При $\alpha \leq 1$ функция параметра потока постепенных отказов после достижения модального значения является монотонно убывающей, тогда как ранее при больших α это был апериодический затухающий процесс с приближением к квазистационарному значению потока равному α/β , причем до достижения модального значения параметр потока отказов соответствует функции плотности распределения [26].

С точки зрения теоретического описания процессов старения произошло замедление индивидуального старения; прежние подходы с эксплуатацией по ресурсу (назначение периода технического обслуживания), эксплуатацией по техническому состоянию (техническое диагностирование свободных каналов) экономически стали неэффективны. Данное утверждение относится и к собственно волоконно-оптическим линиям связи, поскольку старение оптических волокон по показателю изменения затухания показало $\alpha = 0.05$. Соответственно, экономически не обоснованы, не нужны внешние системы контроля и диагностики, необходима встроенная в сеть передачи данных система диагностики [29].

Граничное условие должно рассчитываться для систем, обеспечивающих жизнеобеспечение, или стоимость отказа которых очень велика (например, отдельные категории военной техники). При $\alpha \in (0.01 \dots 1.0)$ для существующего диапазона имеем $\beta_1, \varepsilon_1, t_1, \beta_3, t_3, \varepsilon_2$ имеем всегда выполнение условий экономического преимущества стратегии по мере возникновения отказов для устройств и систем электросвязи сетей общего пользования, не отвечающих за обеспечение безопасности стратегических объектов, остановка которых может причинить вред окружающей среде, жизни и здоровью людей. Особо отметим, что для стратегических объектов может быть определено критическое значение периодичности планового ТО для применения соответствующей стратегии [29]. Стратегии более высокого ранга – с плановым ТО и/или диагностированием – являются, естественно, экономически оправданными для объектов, которые обеспечивают дополнительный эффект, необходимый для компенсации дополнительных затрат на проведение плановых предупредительных замен и/или диагностирование.

Отметим, что внезапные отказы принципиально отличаются от постепенных по причине отсутствия контроля исходных параметров. Параметр потока внезапных отказов для неконтролируемых, непредсказуемых параметров во времени описывается, как и параметр потока постепенных отказов, альфа-распределением. При этом $\omega_{вн}$ в 15–20 раз больше ω_n для полупроводников и всего в три раза – для ламповых изделий. Физически это понятно. Ламповая

техника тем и отличалась – повторяемостью, малым разбросом параметров, переход к полупроводникам, потом – к большим интегральным схемам характеризовался большим разбросом, индивидуальным путем старения каждого изделия.

В конце 1970-х годов еще применялась ламповая техника связи, а во второй половине 1980-х уже применялось оборудование электросвязи на интегральных микросхемах. Сводные результаты исследований [26]: первый этап: 1960–1970-е гг. – электронные лампы с $\alpha \in (2 \dots 15)$; второй этап: 1970–1980-е гг. – дискретные транзисторы и полупроводниковые диоды с $\alpha \in (0.2 \dots 1.0)$, при этом относительный ресурс β увеличился втрое при переходе с ламповой на полупроводниковую основу; третий этап: 1980–1990-е гг. – интегральные схемы (микросхемы), оптические волокна волоконно-оптических линий связи с $\alpha \in (0.01 \dots 0.06)$, при этом относительный ресурс β увеличился еще втрое при переходе на интегральные микросхемы. Фактически качественная смена дважды произошла в течение 30 лет, причем изменение параметра α происходило сравнительно плавно, по мере совершенствования технологии производства РЭА различного назначения и разными изготовителями.

5. Выводы

4.1. Совершенствование технологии изготовления привело не только к существенному увеличению наработки на отказ (ресурса срока службы), но и к необходимости учета индивидуальности изделий. Это стало известно еще в период перехода от ламп к полупроводниковым приборам в 1970-е годы. Фактически это относится и к старению оптических волокон волоконно-оптических линий связи. Нельзя по контролю за одним волокном делать вывод о состоянии прочих волокон в этом же кабеле, нельзя прогнозировать состояние, старение аналогов. Поэтому нужен индивидуальный контроль, а экономически он оправдан исключительно для встроенных систем технической диагностики. Система массового технического обслуживания второй половины XX века трансформировалась в начале XXI века в систему индивидуального встроенного контроля.

4.2. Повышение надежности изделий рассматривается в комплексе с расходами на эксплуатацию. Все многообразие стратегий эксплуатации может быть формализовано и описывается в соответствующих математико-экономических моделях, сопоставимых по критерию минимума расходов/максимума доходов:

4.2.1. С точки зрения формализованного описания (модели, стратегии) для современного оборудования, систем электросвязи практически всегда экономически выигрышной является стратегия по уровню надежности по мере возникновения отказов (СЭ1 по классификации [19]).

4.2.2. Совершенствование, минимизация стоимости встроенной технической диагностики контроля параметров определило большую эффективность стратегии по техническому состоянию (без планового технического обслуживания, СЭ6 по классификации [19]) – для сложных, дорогостоящих систем электросвязи. Отдельно устанавливаемые системы технической диагностики, как правило, не окупаются, экономически не эффективны.

4.2.3. Для систем, обеспечивающих жизнеобеспечение, или особо важных систем электросвязи с очень большой стоимостью отказа (отдельные категории военной техники) существует расчетное предельное значение периодичности планового технического обслуживания. Проведение планового обслуживания экономически оправдано только в этих случаях. При плановом техническом обслуживании всегда эффективно переносить срок обслуживания в случае возникновения внезапного отказа.

4.3. При модернизации технологии производства систем электросвязи произошло два качественных скачка, приведших к увеличению ресурса изделий и увеличению индивидуального разброса скорости изменения параметров, определяющих работоспособность изделия в целом. С точки зрения теории последующие качественные изменения технологии производства не приведут к смене стратегий эксплуатации, так как показатель $\alpha \rightarrow 0$.

Литература

1. Гнеденко Б. В., Беляев Ю. К., Соловьев А. Д. Математические методы в теории надежности. М.: Наука, 1965. 100 с.
2. Герцбах И. Б., Кордонский Х. Б. Модели отказов. М.: Советское радио, 1966. 166 с.
3. Кокс Д. Р., Смит В. А. Теория восстановления. М.: Советское радио, 1967. 300 с.
4. Шишенок Н. А. Основные задачи и особенности эксплуатационно-технического обеспечения проектируемых автоматизированных систем управления // Проблемы надежности систем и средств автоматики. Киев: Техника, 1976. С. 62–68.
5. Смирнов Ю. Г. Прогнозирование эксплуатационно-надежных показателей судовой радиоэлектронной аппаратуры. Л.: Судостроение, 1987. 58 с.
6. Чуев Ю. В., Михайлов Ю. Б. Прогнозирование в военном деле. М.: Воениздат, 1975. 179 с.
7. Зубрилов А. П. Техническая эксплуатация авиационного и радиоэлектронного оборудования. Рига: РКИИГА, 1978. 61 с.
8. Белова Г. Ф., Попов В. Д., Чжо Ко Вин. Сравнение радиационного и термического старения МОП интегральных микросхем // Вопросы атомной науки и техники. Сер. Физика радиационного воздействия на радиоэлектронную аппаратуру. 2012. Вып. 2. С. 29–32.
9. Шульгин Е. И. Физико-математическое описание некоторых технических процессов развития постепенных отказов изделия // Надежность и контроль качества. 1981. № 12. С. 9–15.
10. Алексанян И. Т., Кривошапко В. М. Моделирование параметрических отказов и изучение надежности интегральных схем // Электронная техника. Сер. 8. 1981. № 4. С. 52–58.
11. Проников А. С. Параметрическая надежность машин. М.: Машиностроение, 1978. 592 с.
12. Дружинин Г. В. Надежность систем автоматики. М.: Энергия, 1967. 528 с.
13. Методика расчета надежности изделий с учетом постепенных отказов. М.: Издательство стандартов, 1976. 100 с.
14. Левин Б. Л. Систематизация моделей и стратегий обслуживания технических устройств // Надежность и контроль качества. 1971. № 5. С. 12–65.
15. Барзилович Е. Ю. Модели технического обслуживания сложных систем. М.: Высшая школа, 1982. 231 с.
16. Вайнштейн И. И., Михальченко Г. Е., Вайнштейн Ю. В., Сафонов К. В. О выборе стратегий эксплуатации технических систем // Вестник СибГАУ. 2015. Т. 16, № 3. С. 645–650.
17. Ланецкий Б. Н., Лукьянчук В. В., Фоменко Д. В. Обоснование оптимальных стратегий технического обслуживания и ремонта сложных технических систем // Системы обработки информации. Харьковский университет Воздушных сил им. И. Кожедуба. 2009. Вып. 6 (80). С. 72–74.
18. Володарский В. А. Принципы выбора стратегий обслуживания технических устройств // Современные технологии. Системный анализ. Моделирование. 2012. № 2 (34). С. 110–114.
19. Методика оптимизации периодичности проведения замен технических устройств. М.: Издательство стандартов, 1975. 31 с.
20. Ициксон А. И. Стратегии эксплуатации устройств управления железнодорожными технологическими процессами // Устройства и системы механизации и автоматизации железных дорог. Сб. науч. тр. ВНИИЖТ / под ред. Г. И. Кравченко, А. И. Ициксона. М.: Транспорт. 1988. С. 4–12.
21. Смирнов К. А., Ицкович А. А. Обслуживание и ремонт авиационной техники по состоянию. М.: Транспорт, 1980. 232 с.
22. Шехватов Д. Б. Обслуживание по состоянию. Концепция RCM // Автоматизация в промышленности. 2012. № 9. С. 25–27.
23. Абрамов О. В., Розенбаум А. Н., Климченко В. В. Прогнозирование состояния и планирование эксплуатации систем ответственного назначения // Вестник ДВО РАН. 1996. № 4. С. 65–75.

24. ГОСТ 27.002-89. Надежность в технике. Термины и определения. М.: Издательство стандартов, 1990. 32 с.
25. ГОСТ 18322-78. Система технического обслуживания и ремонта техники. Термины и определения. М.: Стандартинформ, 2007. 7 с.
26. Ициксон А. И. Исследование и оптимизация методов технического обслуживания железнодорожных радиостанций [Текст]: дис. ... канд. техн. наук: 05.22.08: защищена 26.09.89, утв. 21.03.90. Ленинград, 1989. 179 с.
27. ГОСТ 12252-86. Радиостанции с угловой модуляцией сухопутной подвижной службы. Типы, основные параметры, технические требования и методы измерений. М.: Издательство стандартов, 1986. 37 с.
28. ГОСТ 8.207-76. Государственная система обеспечения единства измерений. Прямые измерения с многократными наблюдениями. Методы обработки результатов наблюдений. Основные положения // Приложение 1. Проверка нормальности распределения результатов наблюдений группы / (составной критерий)- согласие опытного (натурного) распределения скорости дрейфа параметров с нормальным. М.: Стандартинформ. 2008. 7 с.
29. Ициксон А. И. Экономическая эффективность технического обслуживания радиоаппаратуры оперативного управления перевозочным процессом/ Автоматика, телемеханика и связь. 1986. № 5. С. 6–8.

*Статья поступила в редакцию 31.01.2019;
переработанный вариант – 21.03.2019.*

Ициксон Александр Исаакович

к.т.н., директор по развитию бизнеса ООО «Компания «Эрланг» (620144, Екатеринбург, ул. Народной воли, д.19а), тел. +7 912 216 7115, e-mail: alex.itsixon@erlang.ru.

Retro analysis of the management operation telecommunication system

Itsixon A.I., Ph.D., Company “Erlang” Ltd

The study focuses on the determination of the strategies of exploitation of systems and telecommunication of the changes of manufacturing technology equipment. Existing research is not linked to the aging of products depending on the technology of manufacture of electronic equipment. The proposed model of correlation maintenance strategy of telecommunication with the improvement of technology. On the basis of the Methodology of the state Committee for standardization identified the main strategy of operation; for mathematics-for-economic models, strategies mapped by the criterion of minimum cost/maximum revenue. Conducted an evolutionary analysis to improve the technology since the 1960's. Conducted field studies of the drift rate (aging) parameters determining the efficiency of devices and telecommunication systems for different elemental components: on the basis of lighting technology, semiconductors, integrated circuits, worked in the 1970s-1990s. The change in the rate of aging was determined. The comparison of operation strategies at the current level of technology development is made. It is proved that there will be no further change in operating strategies. Further improvement of manufacturing technology will be accompanied by quantitative changes.

Keywords: telecommunications, management of maintenance, maintenance strategy, the rate of change of parameters, the criterion of minimum costs of operation.

УДК 004.021

Анализ и оптимизация конвейерных алгоритмов широковещательной передачи стандарта MPI

М. Г. Курносов*

Выполнен теоретический и экспериментальный анализ времени выполнения древовидных алгоритмов операции MPI_Bcast. В модели Хокни определены оптимальные степени деревьев и размеры сегментов в конвейеризированных версиях алгоритмов. Исследование алгоритмов выполнено с учетом деталей их реализации в библиотеки Open MPI. Проведено экспериментальное исследование масштабируемости алгоритмов на вычислительном кластере с сетью связи Gigabit Ethernet.

Ключевые слова: широковещательная передача, трансляционный обмен, broadcast, MPI, параллельное программирование, вычислительные системы.

1. Введение

Со второй половины 1990-х годов основным средством создания параллельных программ, переносимых между широким спектром архитектур вычислительных систем (ВС), являются библиотеки стандарта MPI (message passing interface) [1]. Базовым средством взаимодействия параллельно выполняющихся MPI-процессов является передача сообщений по каналам связи. Стандартом определены три типа информационных обменов между процессами MPI-программы: *двусторонние обмены* (point-to-point communications), *односторонние обмены* (one-sided communications, remote memory access) и *коллективные обмены* (групповые, глобальные, collective communications). В двустороннем обмене участвуют два процесса путем явного обращения к функциям передачи (MPI_Send) и приема сообщения (MPI_Recv). Побочным эффектом этого вида информационного взаимодействия является синхронизация выполнения пары процессов, что может оказать негативное влияние на масштабируемость параллельной программы. При одностороннем обмене в операцию вовлечен только один процесс, который выполняет удаленную операцию записи (MPI_Put) или чтения (MPI_Get) области памяти другого процесса. Синхронизации процессов не происходит, пока соответствующую операцию явно не вызовет один из процессов. Коллективные операции являются самыми ресурсоемкими, в них участвуют все процессы коммуникатора параллельной программы. Стандарт MPI 3.1 определяет 17 коллективных операций, которые можно разделить на три группы: трансляционный обмен (ТО, one-to-all broadcast/scatter), коллекторный (КО, all-to-one gather) и трансляционно-циклический (ТЦО, all-to-all/allgather).

В данной работе основное внимание сосредоточено на алгоритмах реализации операции трансляционной передачи сообщения (one-to-all broadcast) из заданного корневого процесса всем остальным. Данная операция реализуется функцией

MPI_Bcast(buf, count, datatype, root, comm),

которая выполняет передачу из памяти процесса root буфера buf, содержащего count элементов типа datatype, всем процессам коммуникатора comm.

Производительность (время выполнения) операции MPI_Bcast является критически важной для широкого спектра параллельных алгоритмов и программ. В табл. 1 приведены паке-

*Работа выполнена при поддержке РФФИ (проект № 18-07-00624) и частичной финансовой поддержке Президиума РАН (ГЗ 0306-2018-0012).

ты суперкомпьютерного моделирования, производительность которых в значительной степени зависит от времени выполнения операции MPI_Bcast. Результаты получены путем анализа отчетов международного проекта HPC Advisory Council [2, 3]. Функции перечислены в порядке убывания частоты вызовов и суммарного времени их выполнения. Для ряда пакетов приведены значения размеров сообщений, характерные для тестовых наборов данных, на которых выполнялось профилирование пакетов.

Т а б л и ц а 1. Пакеты параллельного моделирования, производительность которых значительно зависит от времени выполнения операции MPI_Bcast

Пакет	Предметная область	Размеры сообщений
Abaqus	Инженерный анализ методом конечных элементов	64-256 байт
ANSYS CFX, FLUENT	Вычислительная аэро-, гидро- и газовая динамика	< 64 Кбайт
FLOW-3D	Вычислительная аэро-, гидро- и газовая динамика	4-16 байт, 1-4 Мбайт
HPCC FFT	Быстрое преобразование Фурье (1D)	
HPCG	Решение систем линейных уравнений с разреженной матрицей	< 2 Кбайт
LAMMPS	Молекулярная динамика	< 256 Кбайт
LS-DYNA	Анализ процессов в задачах механики твердого и жидкого тела	< 4 Кбайт
VASP	Квантово-механическое моделирование	< 256 Кбайт
WRF	Моделирования климатических и погодных явлений	< 16 Кбайт

Существует две основные свободно распространяемые реализации стандарта MPI, Open MPI и MPICH, на базе которых созданы все известные библиотеки MPI, включая коммерческие пакеты ведущих производителей ВС:

- Open MPI: Mellanox HPCX, IBM Spectrum MPI, Fujitsu Post-K MPI, Bull MPI;
- MPICH: MVARICH, Intel MPI, Tianhe2 TH-MPI, Cray MPI, IBM PE MPI.

Для каждой коллективной операции в библиотеке MPI присутствует несколько алгоритмов ее реализации, обладающих разной эффективностью: архитектурно-ориентированные алгоритмы (topology-aware collectives) и универсальные алгоритмы на основе операций MPI_Send/MPI_Recv. Архитектурно-ориентированные алгоритмы имеют высокую эффективность за счет узкой специализации для фиксированной коммуникационной технологии (Mellanox InfiniBand CORE-Direct) или коммуникационной сети ВС (IBM BlueGene/Q, Cray Aries CE, Fujitsu Tofu). Универсальные алгоритмы коллективных операций основаны на использовании только операций MPI_Send и MPI_Recv, которые реализуются нижележащим слоем библиотеки MPI. Такие алгоритмы учитывают только общее число процессов в программе и параметры коллективной операции (размер сообщения, тип данных, номер корневого процесса и др.). Последнее обеспечивает их широкую переносимость между различными архитектурами ВС и сохранение приемлемого для практики уровня производительности.

При вызове коллективной операции библиотека MPI динамически определяет для заданного MPI-коммуникатора (числа процессов) и параметров операции (размера сообщения, типа данных, номера корневого процесса), какой из доступных алгоритмов широковежательной передачи использовать. В текущих реализациях MPICH, MVARICH, Open MPI используются либо фиксированные правила выбора алгоритма на основе размера сообщения, либо выбранный пользователем алгоритм. Значительная часть алгоритмов операции MPI_Bcast имеет настрои-

ваемые параметры, влияющие на общее время выполнения операции.

Цель данной работы – выполнить теоретический и экспериментальный анализ времени выполнения древовидных алгоритмов операции MPI_Bcast и определить оптимальные значения настраиваемых параметров алгоритмов – степени деревьев и размеры сегментов в конвейеризированных версиях алгоритмов. В отличие от ранее выполненных работ [4, 5, 8, 9], в данной работе анализ времени выполнения алгоритмов не ограничивается построением асимптотических оценок, уделено внимание и константам при параметрах α и β используемой модели Хокни. Теоретические оценки подкреплены результатами экспериментов на вычислительной системе с сетью связи стандарта Gigabit Ethernet. Алгоритмы операции MPI_Bcast исследованы с учетом деталей их реализации в библиотеки Open MPI 4.0.0, в разработке которой автор принимал участие.

Организация работы: в следующем разделе дается описание модели Хокни, в рамках которой осуществляется анализ временных характеристик алгоритмов коллективных операций. Далее приводится описание и анализ основных алгоритмов – от простейших к наиболее эффективным. В последних разделах даются рекомендации по использованию алгоритмов.

2. Модель Хокни

Модель Хокни (R. Hockney) – одна из простейших линейных моделей, описывающая вычислительную систему двумя параметрами α и β . Время $t(m)$ передачи сообщения размером m байт между двумя процессами выражается как

$$t(m) = \alpha + m\beta,$$

где α – задержка при передаче сообщения (латентность, latency, startup time), β – время, требуемое для передачи одного байта сообщения. Также данная модель известна под названием «latency-bandwidth model». Параметр α включает в себя как латентность канала связи, так и накладные расходы процессора на передачу/прием сообщения (например, упаковка сообщения в структуры данных протокола MPI, инициализация канала связи). Величина $1/\beta$ отражает пропускную способность канала связи (bandwidth). Как правило, за ее значение принимается верхняя граница пропускной способности для конкретной коммуникационной технологии (например, Gigabit Ethernet, InfiniBand FDR).

На рис. 1 показана иллюстрация модели Хокни. Процесс 0 три раза передает сообщения размером m байт процессу 1. Процесс 1 получает первый байт первого сообщения через α единиц времени и принимает все сообщения за время $m\beta$. Процесс 0 не может начать передачу очередного сообщения до тех пор, пока процесс 1 не получил предыдущее. Поэтому второе сообщение процесс 0 начинает передавать в момент времени $\alpha + m\beta$, а процесс 1 получает его в момент времени $2(\alpha + m\beta)$. Оба процесса заканчивают обмены за время $3(\alpha + m\beta)$.

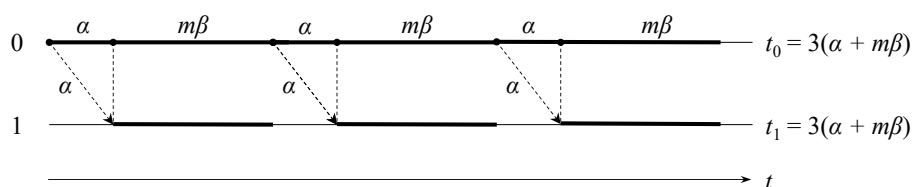


Рис. 1. Модель Хокни: процесс 0 последовательно передает три сообщения размером m байт процессу 1

Модель максимально абстрагируется от архитектуры ВС и деталей реализации инструментария параллельного программирования (игнорируется многопоточность, присутствие неблокирующих коммуникационных операций). Предполагается, что латентность α и пропускная

способность $1/\beta$ не зависят от размера m передаваемого сообщения. Считается, что каналы связи допускают двусторонний обмен данными (bidirectional channels). Коммуникационная сеть является надежной и имеет полносвязную логическую топологию (fully connected). Каждый процесс может выполнять надежные двусторонние обмены операциями MPI_Send и MPI_Recv с любым другим процессом. Системное устройство элементарной машины является однопортовым (single-ported) – каждая элементарная машина ВС (процессор) подключена к сети через один коммуникационный канал. Время передачи информации между любой парой элементарных машин одинаково.

3. Линейный алгоритм

В *линейном алгоритме* (linear broadcast, flat tree) корневой процесс с номером $root$ последовательно передает сообщение процессам $0, 1, \dots, p-1$. На рис. 2 показан псевдокод алгоритма, который выполняют все процессы программы. Алгоритм имеет низкую масштабируемость – время его выполнения линейно зависит от числа p процессов. Из временной диаграммы рис. 3 видно, что наибольшим временем выполнения характеризуется корневой процесс (при любом выборе значения $root$), отсюда время работы линейного алгоритма

$$t = (p - 1)(\alpha + m\beta) = O(p).$$

```

procedure BcastLinear(buf, count, root)
  if rank = root then
    for i = 0 to p- 1 do
      if i != root then
        Send(buf, count, i)
      end if
    end for
  else
    Recv(buf, count, root)
  end if
end procedure

```

Рис. 2. Псевдокод линейного алгоритма Bcast

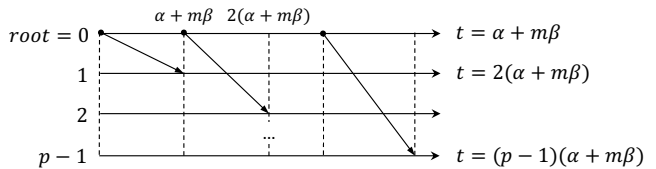


Рис. 3. Временная диаграмма выполнения линейного алгоритма

В линейном алгоритме наименьшим временем t_{min} выполнения характеризуется процесс 0 или 1 (в зависимости от выбора корневого процесса):

$$t_{min} = \alpha + m\beta.$$

Разность между временем выполнения самого медленного процесса $(p - 1)(\alpha + m\beta)$ и самого быстрого $\alpha + m\beta$ линейно зависит от числа процессов, что свидетельствует о значительной разбалансированности времени работы процессов. Последнее негативно сказывается на общем балансе времени выполнения ветвей параллельной программы. Такие алгоритмы не рекомендуется использовать в итерационных вычислительных методах, их использование ограничено случаями, когда другие алгоритмы недоступны по разным причинам (например, не допускают реализацию для заданного размера сообщения или числа процессов). Значительно лучшими динамическими характеристиками обладает класс алгоритмов широковещательной передачи, основанных на логическом выстраивании процессов в деревья различных видов и параллельной передаче сообщений в поддеревьях. Выполним анализ этих алгоритмов и определим их оптимальные параметры в модели Хокни.

4. Алгоритм k -арного дерева

В рассматриваемом алгоритме процессы программы логически выстраиваются в k -арное дерево (k -ary tree). Каждый процесс, зная свой номер $rank$ и общее число p процессов, определяет номер своего родителя и номера не более k дочерних процессов. Основная идея алгоритма – реализовать параллельную передачу сообщения (его ретрансляцию) через промежуточные узлы поддеревьев. Все процессы, за исключением корневого, на первом шаге выполняют прием сообщения от родительского процесса, после этого выполняют последовательную передачу сообщения дочерним процессам, начиная с левого. Например, в бинарном дереве (рис. 4.а) корневой процесс на первом шаге передает сообщение левому дочернему процессу, а на втором шаге – правому (номер шага отмечен над ребром). Одновременно на втором шаге левый дочерний процесс корня передает принятое сообщение своему левому дочернему узлу. На рис. 4.б и рис. 4.в приведены примеры тернарного и k -арного деревьев, а также показаны параметры деревьев и оценки сверху времени t выполнения алгоритмов.

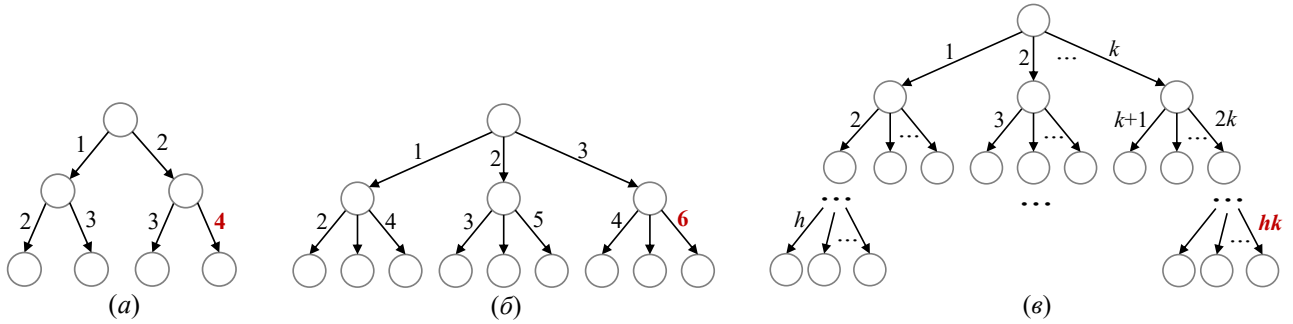


Рис. 4. Структура совершенных k -арных деревьев ширококестельной передачи:

- а – бинарное дерево (binary tree, $k = 2, p = 7, h = 2$): $t = 2 \lfloor \log_2 p \rfloor (\alpha + m\beta)$;
- б – тернарное дерево (ternary tree, $k = 3, p = 13, h = 2$): $t = 3 \lfloor \log_3 p \rfloor (\alpha + m\beta)$;
- в – k -арное дерево (k -ary tree): $t = k \lfloor \log_k p \rfloor (\alpha + m\beta)$

Заметим, что поиск процессом своего положения в дереве в худшем случае требует порядка $O(\log p)$ арифметико-логических операций. Как правило, подобные вычисления параметров деревьев выполняются только на этапе создания MPI-коммуникаторов и сохраняются в памяти каждого процесса.

Оценим время выполнения алгоритма в модели Хокни. Для заданного числа p процессов k -арное дерево может быть построено различными способами: иметь вид завершеного k -арного дерева (complete k -ary tree) либо k -арного дерева, сбалансированного по числу процессов в поддеревьях. Последняя форма дерева наряду с другими используется в библиотеке Open MPI (coll/base) для реализации операций MPI_Bcast и MPI_Reduce. Указанные деревья отличаются схемой заполнения листьями последнего уровня. В завершеном дереве узлы добавляются слева направо, все внутренние вершины имеют степень k (рис. 5.а). В сбалансированном по числу узлов дереве вершины добавляются слева направо, но к родителям из разных поддеревьев. Например, на рис. 5.б листовые узлы равномерно распределены по поддеревьям с корнями в узлах 1, 2 и 3.

Если дерево из p процессов имеет вид совершенного k -арного дерева (perfect k -ary tree), в котором все внутренние узлы имеют степень k , а листья находятся на одном уровне h , то время t_{max} выполнения алгоритма определяется временем работы самого правого листового процесса (рис. 4.в):

$$t_{max} = kh(\alpha + m\beta) = k \lfloor (\log_k(p(k-1) + 1) - 1) \rfloor (\alpha + m\beta).$$



Рис. 5. Тернарное дерево широковещательной передачи ($p = 11, root = 0$):

a – завершенное тернарное дерево;

b – тернарное дерево, сбалансированное по числу узлов в поддеревьях

Наименьшим временем выполнения характеризуется самый левый лист

$$t_{min} = h(\alpha + m\beta) = \lceil (\log_k(p(k-1) + 1) - 1) \rceil (\alpha + m\beta).$$

Если процессов не хватает для формирования совершенного вида дерева, то время t_{max} выполнения может варьироваться в следующем интервале:

$$k(h-1)(\alpha + m\beta) \leq t_{max} \leq kh(\alpha + m\beta).$$

Например, на рис. 4 для формирования совершенного дерева не хватает двух процессов, здесь наибольшее время выполнения имеет процесс 9. Далее, не теряя общности, будем полагать, что процессы образуют совершенное дерево. Указанное допущение не влияет на дальнейшие рассуждения, так как при анализе времени выполнения алгоритма нас интересуют оценки сверху времени выполнения. Тогда

$$t = k \lceil \log_k p \rceil (\alpha + m\beta) = O(\log p).$$

На практике при реализации алгоритмов и разработке библиотек MPI встает вопрос о выборе оптимального значения степени k дерева для заданного числа p процессов. Такое решение принимается динамически при формировании MPI-коммуникаторов. В библиотеках Open MPI и MVAPICH по умолчанию используется значение $k = 4$ для любого числа процессов. Определим оптимальное в модели Хокни значение k . Запишем оценку сверху времени выполнения алгоритма как функцию от параметра k и найдем его значение, при котором время выполнения алгоритма минимально:

$$t(k) = k \log_k p (\alpha + m\beta), \quad \frac{\partial t(k)}{\partial k} = \frac{\ln k - 1}{\ln^2 k} = 0, \quad k = e \approx 2.7183.$$

Таким образом, наименьшим в модели Хокни среди k -арных деревьев временем выполнения обладает тернарное дерево ($k = 3$). Полученная оценка хорошо согласуется с практикой – минимальное время выполнения достигается на k -арных деревьях небольших степеней. С увеличением значения k наблюдается рост времени выполнения алгоритма. На рис. 5 показаны результаты измерения времени выполнения алгоритма k -арного дерева широковещательной рассылки для различных значений k . Измерения выполнены для реализации алгоритма в библиотеке Open MPI 4.0.0 общепринятым тестом Intel MPI Benchmarks (приведено среднее время выполнения по не менее 40 запускам для каждого k) для 128 процессов: 16 двухпроцессорных узлов (2 x Intel Xeon Quad), сеть связи – Gigabit Ethernet.

В отличие от линейного, рассматриваемый алгоритм хорошо масштабируется – время его работы логарифмически зависит от числа процессов. В то же время алгоритм имеет существенный недостаток – фиксированная степень k корня и внутренних узлов является причиной их недостаточной загруженности операциями передачи. Так, в тернарном дереве (рис. 4.б)

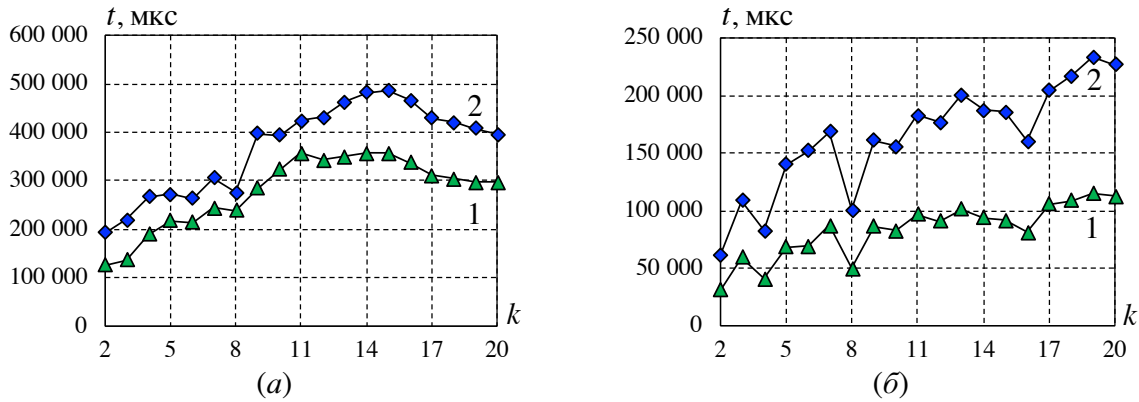


Рис. 6. Зависимость времени выполнения алгоритма от степени k дерева
 ($p = 128$, 1 – размер сообщения 1 МБ; 2 – размер сообщения 2 МБ):
 а – алгоритм k -арного дерева;
 б – алгоритм k -номиального дерева

после выполнения трех операций корень завершает свою работу, хотя в это время в дереве присутствуют узлы, ожидающие сообщения. Указанного недостатка лишен алгоритм k -номиального дерева.

5. Алгоритм k -номиального дерева

В алгоритме k -номиального дерева процессы логически выстраиваются в дерево обменов, представляющее совокупность k -номиальных деревьев (k -nomial tree). При $k = 2$ дерево называется *биномиальным* (binomial tree), а при $k = 3$ – *триномиальным* (trinomial tree). На рис. 7 показаны примеры биномиального, 3-номиального и 4-номиального деревьев для широковещательной передачи сообщения из корневого процесса 0. Каждый процесс, за исключением корневого, выполняет два шага – прием сообщения от родителя и его передачу дочерним процессам, начиная с самого левого. Номера процессов трактуются как числа, записанные в системе счисления с основанием k . Номер родительского процесса вычисляется путем сброса в номере текущего процесса самого младшего разряда с положительным весом (least significant digit, LSD). Например, в биномиальном дереве номер родителя для процесса $12 = 1100_2$ равен $8 = 1000_2$, который получен путем сброса в числе 12 третьего бита (младшего значащего разряда). Количество дочерних узлов любого процесса не превышает $(k - 1)\lceil \log_k p \rceil$. Таким образом, в отличие от k -арных деревьев, в k -номиальных степень корня (внутреннего узла) не фиксирована, а логарифмически зависит от числа вершин в его поддереве. Последнее позволяет k -номиальным деревьям эффективнее загрузить корень и внутренние процессы передачей сообщений. Из примера на рис. 7 видно, что наименьшим временем выполнения характеризуется биномиальное дерево, которому требуется 4 шага для завершения операции.

В общем случае время $t(k)$ выполнения алгоритма определяется временем работы корневого процесса – временем завершения передачи сообщения самому правому дочернему узлу

$$t(k) = (k - 1)\lceil \log_k p \rceil(\alpha + m\beta).$$

Монотонное возрастание функции $t(k)$ позволяет сделать вывод об оптимальности биномиальных деревьев в модели Хокни. Результаты экспериментов на вычислительных кластерах подтверждают оптимальный характер биномиальных деревьев. На рис. 6.б показано монотонное возрастание времени выполнения алгоритма k -номиального дерева с увеличением степени дерева на вычислительном кластере с сетью связи Gigabit Ethernet.

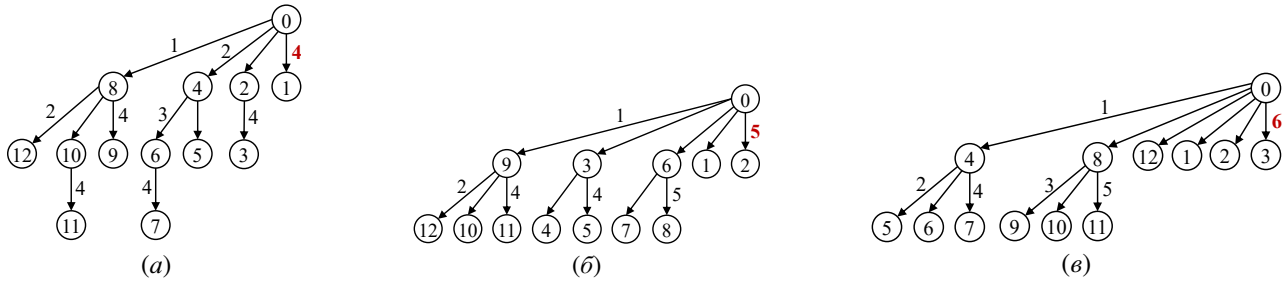


Рис. 7. Структура k -номиальных деревьев широковещательной передачи:

- a – биномиальное дерево (binomial tree, $k = 2, p = 13, h = 3$);
 b – триниомальное дерево (trinomial tree, $k = 3, p = 13, h = 2$);
 v – 4-номиальное дерево (4-nomial tree, $k = 4, p = 13, h = 2$)

Несмотря на указанные достоинства, данный алгоритм не учитывает возможность одновременного приема и передачи сообщений одним процессом по дуплексным каналам связи, которые характерны для подавляющего большинства коммуникационных сетей ВС (Gigabit Ethernet, InfiniBand). Возможным решением является разработка алгоритма, в котором исходное сообщение разбивается на сегменты и передается по конвейеру, так, чтобы промежуточные процессы одновременно передавали текущий сегмент и принимали следующий. Рассмотрим этот подход на примере анализа и оптимизации конвейерного алгоритма для больших сообщений.

6. Конвейерный алгоритм

В конвейерном алгоритме (linear pipeline) исходное сообщение длиной m байт разбивается на $\lceil m/s \rceil$ сегментов по s байт, которые по цепочке передаются от процесса к процессу. Промежуточные процессы после получения очередного сегмента m_i начинают использовать дуплексный режим передачи – передавать текущий сегмент m_i и одновременно с этим принимать следующий сегмент m_{i+1} . На рис. 8 приведена диаграмма выполнения конвейерного алгоритма для четырех процессов и сообщения, разбитого на три сегмента m_1, m_2 и m_3 . Промежуточные процессы 1 и 2 конвейера после получения сегмента m_1 сразу начинают его передавать следующему процессу и одновременно с этим принимать сегмент m_2 .

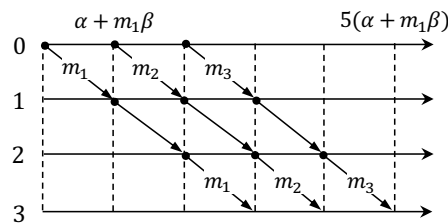


Рис. 8. Диаграмма выполнения конвейерного алгоритма ($p = 4$, три сегмента)

Наибольшим временем выполнения характеризуется процесс $p - 1$, который последним получает заключительный сегмент сообщения. Первый сегмент он получает в момент времени $(p - 1)(\alpha + s\beta)$, второй в $p(\alpha + s\beta)$, третий в $(p + 1)(\alpha + s\beta)$ и т.д. Заключительный сегмент процесс $p - 1$ получит в момент времени $(p - 2 + \lceil \frac{m}{s} \rceil)(\alpha + s\beta)$. Таким образом, время работы конвейерного алгоритма

$$t(s) = \left(p - 2 + \left\lceil \frac{m}{s} \right\rceil \right) (\alpha + s\beta).$$

Определим оптимальный размер s_{opt} сегмента, при котором алгоритм характеризуется минимальным в модели Хокни временем выполнения:

$$\frac{\partial t(s)}{\partial s} = -\frac{m\alpha}{s^2} + (p-2)\beta = 0, \quad s_{opt} = \sqrt{\frac{m\alpha}{(p-2)\beta}}.$$

На рис. 9.а показана зависимость времени выполнения конвейерного алгоритма в модели Хокни от размера сегмента (параметры $\alpha = 5 \cdot 10^{-5}$ с, $\beta = 4.7 \cdot 10^{-8}$ с получены тестом Intel MPI Benchmarks для сети связи Gigabit Ethernet). Красными точками показаны оптимальные размеры s_{opt} сегментов. Полученная теоретическая оценка s_{opt} дает хорошее приближение к размеру сегмента, эффективного для реальных коммуникационных сетей. Ее использование позволяет сократить границы поиска эффективного размера сегмента для конкретной коммуникационной технологии (Gigabit Ethernet, InfiniBand). На рис. 9.б показаны экспериментальные результаты – время выполнения алгоритма на вычислительном кластере с сетью связи Gigabit Ethernet и установленные границы (суб)оптимальных размеров сегментов [1536, 3072] (выделены цветом).

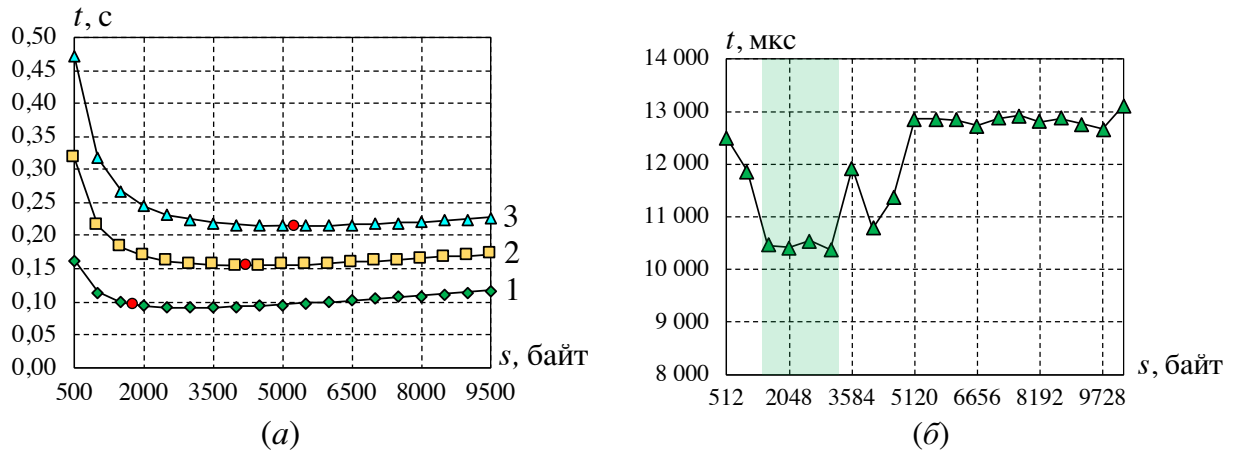


Рис. 9. Зависимость времени выполнения конвейерного алгоритма от размера s сегмента (128 процессов):

а – модели Хокни ($\alpha = 5 \cdot 10^{-5}$ с, $\beta = 4.7 \cdot 10^{-8}$ с): 1 – $m = 1$ МБ; 2 – $m = 2$ МБ; 3 – $m = 3$ МБ;

б – эксперимент на кластере (Open MPI 4.0.0, тест Intel MPI Benchmarks, сеть Gigabit Ethernet): $m = 1$ МБ

На рис. 10 показана зависимость в модели Хокни времени выполнения алгоритма биномиального дерева и конвейерного алгоритма с оптимальным размером сегмента. Для сообщений небольших размеров эффективными являются алгоритмы бинарного и биномиального деревьев. Для сообщений значительных размеров предпочтительным является конвейерный алгоритм. Очевидно, что для сообщений средних размеров разумно применить комбинированный подход – древовидный алгоритм с конвейеризацией передачи сообщений.

7. Алгоритм бинарного дерева с конвейеризацией

В алгоритме *бинарного дерева с конвейеризацией передачи сообщений* (pipelined binary tree) процессы логически организованы в бинарное дерево. Исходное сообщение длиной m байт разбивается на $\lceil m/s \rceil$ сегментов по s байт, которые передаются по дереву с учетом дуплексного режима работы каналов связи. На рис. 11 показан пример широковещательной

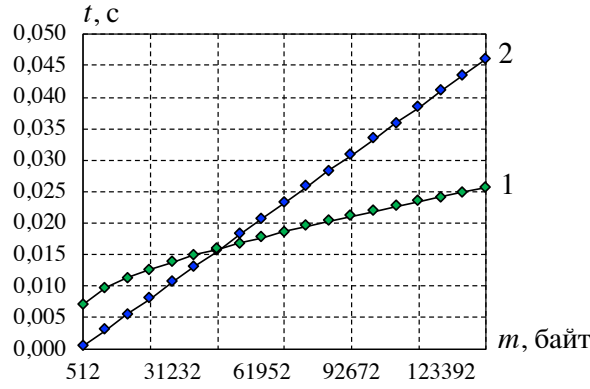


Рис. 10. Зависимость времени выполнения алгоритмов операции Bcast в модели Хокни от размера m сообщения ($\alpha = 5 \cdot 10^{-5}$ с, $\beta = 4.7 \cdot 10^{-8}$ с):

1 – конвейерный алгоритм с оптимальным размером сегмента; 2 – биномиальное дерево

передачи сообщения, разбитого на 3 сегмента. Временные шаги передачи отдельных сегментов отмечены на дугах. На шаге 3 процесс 1 начинает передачу первого сегмента процессу 4 и одновременно принимает второй сегмент от процесса 0. В таблице справа (рис. 11) подчеркнуты пары процессов, которые одновременно принимают и передают сегмент сообщения – используют дуплексный режим работы каналов связи. Как видно, не все процессы и не на каждом временном шаге в полной мере используют доступную полосу пропускания каналов связи. По этой причине алгоритму бинарного дерева с конвейеризацией передачи сообщений не удается достигнуть полной загрузки дуплексных каналов связи (full bandwidth).

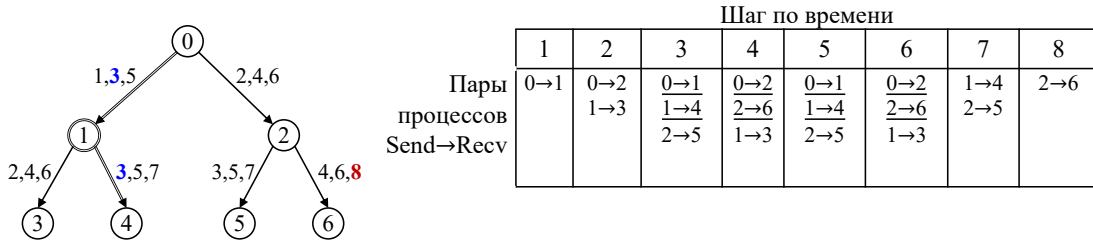


Рис. 11. Бинарное дерево с конвейеризацией передачи сообщений ($p = 7$, $root = 0$, 3 сегмента, числа над дугами – временные шаги передачи сегментов)

Оценим сверху время выполнения алгоритма. Как было показано ранее для k -арных деревьев, последним первый сегмент принимает правый листовой процесс. Следующий сегмент он примет через число шагов, равное степени дерева, и т.д. Таким образом, время $t(s)$ широковещательной передачи алгоритмом бинарного дерева с конвейеризацией сообщений имеет вид

$$t(s) = 2 \left(\lceil \log_2 p \rceil + \frac{m}{s} - 1 \right) (\alpha + s\beta).$$

Рассуждая аналогично, можно получить оценку и для биномиального дерева с конвейеризацией сообщений

$$t(s) = \left(\lceil \log_2 p \rceil + \frac{m}{s} - 1 \right) (\alpha + s\beta).$$

Определим оптимальный размер сегмента для алгоритма бинарного дерева с конвейеризацией передачи сообщений:

$$\frac{\partial t(s)}{\partial s} = -\frac{2m}{s^2} \alpha + 2(\log_2 p - 1)\beta = 0, \quad s_{opt} = \sqrt{\frac{m\alpha}{(\log_2 p - 1)\beta}}.$$

На рис. 12 показана зависимость времени выполнения алгоритма бинарного дерева с конвейеризацией передачи сообщений от размера s сегмента. Как видно из экспериментальных результатов, наименьшее время выполнения достигнуто при размере сегмента 13 312 байт. Теоретически установленный оптимальный размер сегмента $s_{opt} = 13\,632$ байт. В целом полученное значение дает хорошее приближение для выбора субоптимального значения сегмента на практике.

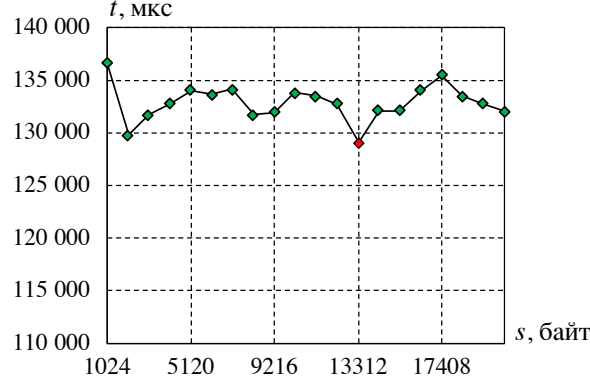


Рис. 12. Зависимость времени выполнения алгоритма бинарного дерева с конвейеризацией сообщений от размера s сегмента ($p = 128$, $m = 1$ МБ, Open MPI 4.0.0, тест Intel MPI Benchmarks, сеть Gigabit Ethernet)

8. Обсуждение результатов и рекомендации

Известна [9] теоретическая нижняя граница времени t_{Bcast} выполнения широковещательной передачи в модели Хокни

$$t_{Bcast} \geq \min\{\lceil \log_2 p \rceil \alpha, m\beta\}.$$

Оценим отклонение времени работы рассмотренных алгоритмов от указанной нижней границы.

Для сообщений небольших размеров ($\alpha > m\beta$) время работы алгоритма биномиального дерева определяется коэффициентом $\lceil \log_2 p \rceil$ при параметре α – вкладом латентности канала связи в суммарное время работы алгоритма

$$t_{Binomial} = \lceil \log_2 p \rceil \alpha + m \lceil \log_2 p \rceil \beta.$$

Для рассматриваемого случая сообщений небольших размеров значение коэффициента при параметре α с точностью до константы совпадает со значением нижней границы t_{Bcast} , что свидетельствует об асимптотической оптимальности алгоритма биномиального дерева для сообщений небольших размеров. При сообщениях средних и больших размеров ($\alpha < m\beta$), когда вклад латентности канала связи не оказывает решающего значения, алгоритм биномиального дерева имеет при параметре β коэффициент в $\log_2 p$ раз больший значения нижней границы. На рис. 13 приведен пример, иллюстрирующий эти два случая.

Линейная зависимость времени выполнения конвейерного алгоритма (linear pipeline) ограничивает его применение случаем незначительного числа процессов, а также сообщений больших размеров:

$$t_{Pipeline} = \left(p - 2 + \left\lceil \frac{m}{s} \right\rceil \right) (\alpha + s\beta).$$

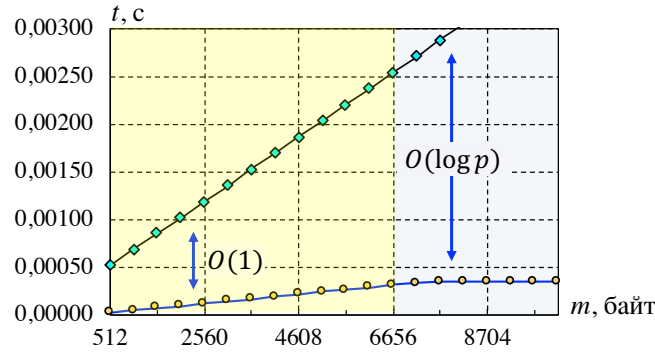


Рис. 13. Отклонение времени выполнения алгоритма биномиального дерева от нижней границы времени операции Bcast ($\alpha = 5 \cdot 10^{-5}$ с, $\beta = 4.7 \cdot 10^{-8}$ с, $p = 128$, $root = 0$)

Латентность алгоритма бинарного дерева с конвейеризацией сообщений (pipelined binary tree) в $2 + \frac{m/s}{\log_2 p}$ раз превышает значение латентности нижней границы

$$t_{\text{PipelinedBinTree}} = 2 \left(\lceil \log_2 p \rceil + \left\lceil \frac{m}{s} \right\rceil - 1 \right) (\alpha + s\beta).$$

Нетрудно заметить, что при небольших размерах сообщений и сегментов отношение времени работы алгоритма к нижней границе равно константе, что свидетельствует об асимптотической оптимальности алгоритма для данного случая. Пропускная способность алгоритма в $2 + \frac{2s \log_2 p}{m}$ раз превышает значение пропускной способности нижней границы. При больших размерах сообщений отношение равно константе, что говорит об асимптотической оптимальности алгоритма по пропускной способности.

На рис. 14 в модели Хокни показаны диапазоны размеров сообщений и рекомендуемые к использованию на них алгоритмы широкоэвещательной передачи. Для сообщений небольших размеров рекомендуется использовать алгоритм биномиального дерева. Для сообщений средних размеров уже целесообразно задействовать конвейеризацию передачи сообщений и алгоритм бинарного или биномиального дерева. Для сообщений больших размеров и небольшого числа процессов рекомендуется использовать конвейерный алгоритм.

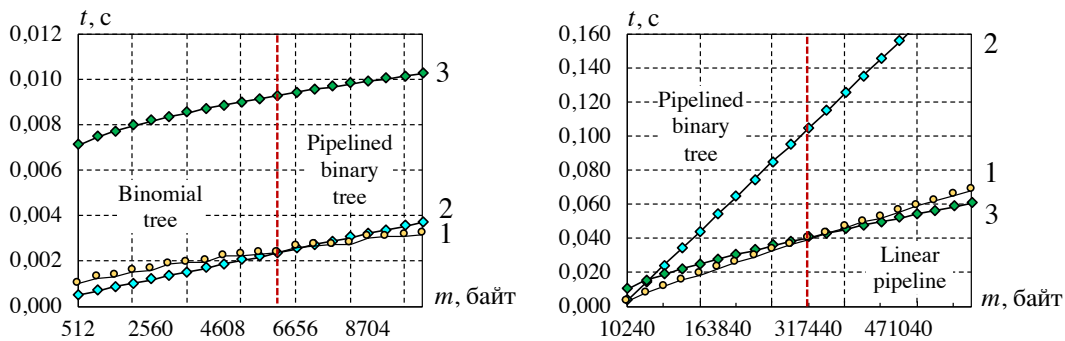


Рис. 14. Зависимость времени выполнения алгоритмов MPI_Bcast в модели Хокни от размера m сообщения ($\alpha = 5 \cdot 10^{-5}$ с, $\beta = 4.7 \cdot 10^{-8}$ с, $p = 128$, $root = 0$):

- 1 – алгоритм бинарного дерева с конвейеризацией (с оптимальным размером сегмента);
- 2 – алгоритм биномиального дерева;
- 3 – конвейерный алгоритм с оптимальным размером сегмента

9. Заключение

Выполненный теоретический и экспериментальный анализ времени выполнения древовидных алгоритмов операции MPI_Bcast показал большую эффективность алгоритма биномиального дерева по сравнению с алгоритмами тернарного и бинарного деревьев. Для алгоритмов с сегментацией сообщений установлена эффективность алгоритмов бинарного и биномиального деревьев для сообщений средних размеров. Для небольшого числа процессов и сообщений значительных размеров эффективным является конвейерный алгоритм.

Построенные аналитические выражения носят рекомендательный характер и позволяют сузить на практике поиск эффективных степеней деревьев и размеров сегментов для алгоритмов с конвейеризацией передачи сообщений.

Литература

1. MPI-3.1 Standard [Электронный ресурс]. URL: <https://www.mpi-forum.org/docs/mpi-3.1/mpi31-report.pdf> (дата обращения: 16.03.2019).
2. HPC Advisory Council Best Practices [Электронный ресурс]. URL: http://hpcadvisorycouncil.com/best_practices.php (дата обращения: 16.03.2019).
3. Dongarra J., Beckman P., Moore T. International Exascale Software Project Roadmap // The International Journal of High Performance Computing Applications. 2011. V. 25. P. 3–60.
4. Thakur R., Rabenseifner R., Gropp W. Optimization of collective communication operations in MPICH // Int. Journal of High Performance Computing Applications. 2005. V. 19 (1). P. 49–66.
5. Bruck J. et al. Efficient Algorithms for All-to-All Communications in Multiport Message Passing Systems // IEEE Trans. Parallel Distrib. Syst. 1997. V. 8 (11). P. 1143–1156.
6. Курносов М. Г. Алгоритмы трансляционно-циклических информационных обменов в иерархических распределенных вычислительных системах // Вестник компьютерных и информационных технологий. 2011. № 5. С. 27–34.
7. Balaji P., Buntinas D., Goodell D. et al. MPI on Millions of Cores // Parallel Processing Letters. 2011. V. 21 (1). P. 45–60.
8. Hoefler T., Moor D. Energy, Memory, and Runtime Tradeoffs for Implementing Collective Communication Operations // Journal of Supercomputing Frontiers and Innovations. 2014. V. 1 (2). P. 58–75.
9. Sanders P., Speck J., Traff J. L. Two-Tree Algorithms for Full Bandwidth Broadcast, Reduction and Scan // Parallel Computing. 2009. Vol. 35 (12). P. 581–594.

Статья поступила в редакцию 20.03.2019.

Курносов Михаил Георгиевич

д.т.н., профессор кафедры вычислительных систем СибГУТИ (630102, Новосибирск, ул. Кирова, 86), тел. (383) 269-82-86, e-mail: mkurnosov@sibgti.ru; старший научный сотрудник Института физики полупроводников им. А. В. Ржанова СО РАН (630090, Новосибирск, пр-т. Лаврентьева, 13), e-mail: mkurnosov@isp.nsc.ru.

Analysis and optimization of pipelined broadcast algorithms**M. Kurnosov**

Theoretical and experimental analysis of MPI_Bcast algorithms are performed. The optimal tree degrees and segment sizes for pipelined versions of algorithms are obtained. Algorithms were investigated according to their implementation in the Open MPI. Theoretical results are consistent with experiments on a computer cluster with a Gigabit Ethernet communication network.

Keywords: broadcast, MPI, parallel programming, high-performance computing.

УДК: 519.711.3+519.715+621.395.72

Теоретико-множественная модель для решения задачи системного анализа и синтеза в телекоммуникациях

В. Р. Губкина

Рассматривается теоретико-множественная модель для решения задачи системного анализа и синтеза в сложной нелинейной системе с учетом входных параметров, требований к выходным характеристикам, вариативности построения архитектур системы, внешних деструктивных воздействий на нее. Приведены результаты апробации на примере опорной цифровой автоматической телефонной станции.

Ключевые слова: сложная нелинейная динамическая система, системный анализ и синтез, теоретико-множественная модель, электромагнитная совместимость, цифровая автоматическая телефонная станция.

1. Введение

Широкое распространение цифровых устройств в современном телекоммуникационном оборудовании, применение компьютерных сетей при значительных информационных потоках приводит к возрастанию вероятности отказов в телекоммуникационных системах при воздействии электромагнитного импульса и электромагнитных факторов природного и техногенного происхождения (разрядов молнии, разрядов статического электричества, высоковольтных линий электропередачи, электромагнитных полей радиопередающих и радиоприемных станций). Поэтому в этих условиях становятся актуальными проблемы защиты и обеспечения функциональной безопасности информационных и телекоммуникационных систем.

Последствиями в случае отказа технических средств из-за воздействия электромагнитных помех могут быть как существенные материальные потери, так и угроза безопасности и жизни пользователя. Исследование вопросов устойчивости телекоммуникационного оборудования к электромагнитным помехам является необходимым и актуальным аспектом развития современной техники. В условиях нарастающего потребления электроэнергии требования к качеству электроэнергии как к фактору, во многом определяющему безотказное функционирование телекоммуникационного оборудования, должны соответствовать ГОСТ Р 54149-2010 и ряду других стандартов. И хотя в последние годы улучшилась ситуация по исследованию сетевых помех, задачи устойчивости телекоммуникационного оборудования к их воздействию не могут иметь оптимального решения, основываясь только на существовании нормативных требований. Необходимо проведение исследований опасных видов помех и предотвращение нежелательных от них последствий. В частном случае это защита телекоммуникационного оборудования от высоковольтных импульсных помех, представляющих собой один из наиболее опасных видов сетевых помех.

Проблемам оценки уровней кондуктивных помех и защиты телекоммуникационного оборудования от них посвящено значительное количество монографий и публикаций как в отечественной, так и в зарубежной литературе. Так, М. А. Горловым, А. С. Кривовым, В. Ю. Кириловым, Л. Н. Кичиевым, В. Ю. Рогинским, М. В. Марченко, А. С. Строгоновым и другими авторами были рассмотрены физические основы возникновения импульсных и динамических

помех, механизмы их воздействия на электронную аппаратуру, методы и средства ее защиты на стадиях схмотехнического и конструктивного проектирования.

В современные телекоммуникационные системы внедрены информационно-вычислительные комплексы повышенного быстродействия, локальные и глобальные вычислительные сети. Построение систем телекоммуникаций на основе компьютерных сетей позволяет расширить ряд услуг, внедрить новые идентификационные документы в системы МВД, миграционной службы и т.п. При решении данной проблемы приходится сталкиваться с рядом трудностей, которые определяются несовершенством инфраструктуры зданий и помещений, предназначенных для размещения аппаратуры нового поколения, отсутствием ряда стандартов и методик проектирования. В первую очередь это относится к задачам информационной безопасности и электромагнитной совместимости (ЭМС), так как устаревшее оборудование было в меньшей степени восприимчиво ко многим типам разного вида помех. Если решать проблему ЭМС на этапе отладки установленных телекоммуникационных систем, это приведёт к значительным затратам и с высокой долей вероятности потребует реконструкции помещений и переналадки оборудования. Поэтому очевидным является решение задачи ЭМС телекоммуникационного оборудования на более ранних этапах их внедрения, которые включают в себя строительство зданий, отделку помещения, размещение оборудования, прокладку кабельных систем, силовых цепей, выполнение заземления и т.п.

Многообразие архитектур телекоммуникационного оборудования с источниками вторичного электропитания, схмотехнических решений преобразователей, входящих в их состав, осложнили проблему ЭМС и одновариантности для достижения оптимального построения телекоммуникационной системы. При современном уровне развития цифрового оборудования связи стоит задача проверки функционирования телекоммуникационных систем при деструктивных воздействиях для удовлетворения требований по надежности системы и ЭМС. Многообразие компонентов сложной системы, существующие нелинейные взаимосвязи и большая вариативность являются препятствием к поиску простого решения данной задачи. В этом случае находит место применение методов системного анализа, который позволяет устанавливать и структурировать связи между элементами сложных нелинейных систем.

2. Постановка задачи

Современное телекоммуникационное оборудование и системы управления обладают большим количеством элементов, множеством связей и взаимосвязей, а также значительным объемом обрабатываемой информации. Для таких систем, включающих в себя большое число взаимосвязанных подсистем, более эффективным является для начала наметить основные подсистемы и определить главные связи между ними, после чего переходить к более детальному моделированию каждой из подсистем. На начальном этапе проектирования характерна ограниченность информации о свойствах будущей системы, поэтому сначала требуется обратиться к структуре системы и к информации, которую она содержит. Изучение особенностей этой информации относится к морфологическому анализу системы.

Морфологическое описание нужно для представления строения системы. Уровень детализации и глубина описания определяют, какие компоненты системы будут рассмотрены в качестве элементов. Морфологическое описание строится по иерархичному принципу. Уровень конфигурации морфологии определяется таким образом, чтобы создать представление об основных свойствах системы [1, 2].

Фундаментальный вклад в решение задач анализа и синтеза сложных динамических систем внесли российские и зарубежные ученые: С. В. Акимов, Д. А. Аветисян, В. Н. Волкова, А. А. Денисов, Д. Н. Колесников, Н. Н. Моисеев, В. М. Одрин, Ф. Цвикки и др. В работах ученых используется морфологический подход к решению задач системного анализа и синтеза, который был изложен Ф. Цвикки в 30-е годы и в дальнейшем получил развитие

С. В. Акимовым и В. М. Одриным. Метод позволяет сформировать упорядоченное морфологическое множество на уровне идентификации и представить сложную систему в виде морфологического дерева или морфологической таблицы [3, 4]. Результаты этих исследований сыграли огромную роль в формировании основных представлений о динамических свойствах сложных систем. Вместе с тем решения задач анализа и синтеза сложных систем носили разрозненный характер, не имели объединяющей концепции для интеграции методик исследований [5, 6]. Успехи, достигнутые в применении вычислительных методов, затенили теоретические основополагающие аспекты проблемы.

Задача системного анализа и синтеза сложных нелинейных систем с многовариантностью их построения является достаточно сложной и мало разработанной. Для нахождения оптимального решения существует ряд препятствий: неоднородность данных описания объекта, большая размерность обрабатываемой информации, усложняющая переработку данных, а также большое количество воздействующих факторов. Применение теоретико-множественных представлений позволит отобразить сложную многоуровневую систему и протекающие в ней процессы формальным языком при помощи совокупности множеств. Использование метода морфологического анализа и синтеза позволит определить классы объектов исследуемой области, выделить их классификационные признаки для формирования множеств из однородных элементов. Интеграция теоретико-множественных представлений, структурного и параметрического методов синтеза на базе морфологического множества иерархической структуры сложной нелинейной системы позволит разработать критерии и модели для синтеза оптимальной структуры с высокой надежностью функционирования.

Данная задача рассмотрена на примере цифровой автоматической телефонной станции с источником вторичного электропитания и соединительными линиями. Поставлены задачи по повышению эффективности принятия решений в сложной динамической структуре: выбор классификационных признаков, математическое описание объекта, создание имитационных моделей, разработка методов и алгоритмов структурно-параметрического синтеза. Для оценки функционирования сложной нелинейной системы с замкнутой обратной связью как в статических, так и в динамических режимах работы при большом многообразии внешних деструктивных воздействиях и смене состояний разработаны математические модели с использованием теоретико-множественного языка описания.

3. Теоретико-множественная модель

Представим сложную нелинейную систему совокупностью множеств, отображающих описание системы и протекающие в ней процессы (рис. 1):

- множество параметров системы задают начальные условия, требуемые для построения сложной системы $P = \{p_h, h \in G\}, G = \{1, \dots, g\}$;
- множество требований сложной системы $TR = \{tr_r, r \in U\}, U = \{1, \dots, r\}$ задают требования к системе и участвуют в формировании множества структур нелинейной системы $S = \{s_c, c \in MC\}, MC = \{1, \dots, mc\}$;
- множество элементов системы $L = \{l_e, e \in ME\}, ME = \{1, \dots, me\}$ описывают компоненты, входящие в состав системы, которые отбираются в зависимости от структуры системы;
- множество алгоритмов расчета статистических режимов работы формируются множеством $A = \{a_{al}, al \in AL\}, AL = \{1, \dots, al\}$;
- множество деструктивных воздействий $D_i(t_i) = \{d_i, i \in N\}, N = \{1, \dots, n\}$ описывает различные виды внешних деструктивных воздействий и является исходными данными для расчета моделей в динамических режимах работы;

- множество выходных сигналов $K_j(t_j) = \{k_j, j \in P\}$, $P = \{1, \dots, p\}$ используется для анализа множества входных $IM = \{IM_1, IM_2, \dots, IM_{N_{IM}}\}$ и синтеза множества выходных $OM = \{OM_1, OM_2, \dots, OM_{N_{OM}}\}$ модальностей.

Множество моделей в динамических режимах работы (МД) обеспечивают проверку системы на устойчивость к внешним деструктивным воздействиям D , где сложная система представляется в виде схемы с сосредоточенными параметрами и включает пассивные элементы и ключи для перехода из одного состояния в другое. Воздействие описывается функцией $D_n(t_n)$ и поступает на вход k -го блока, действие которого описывается функцией R (рис. 2).

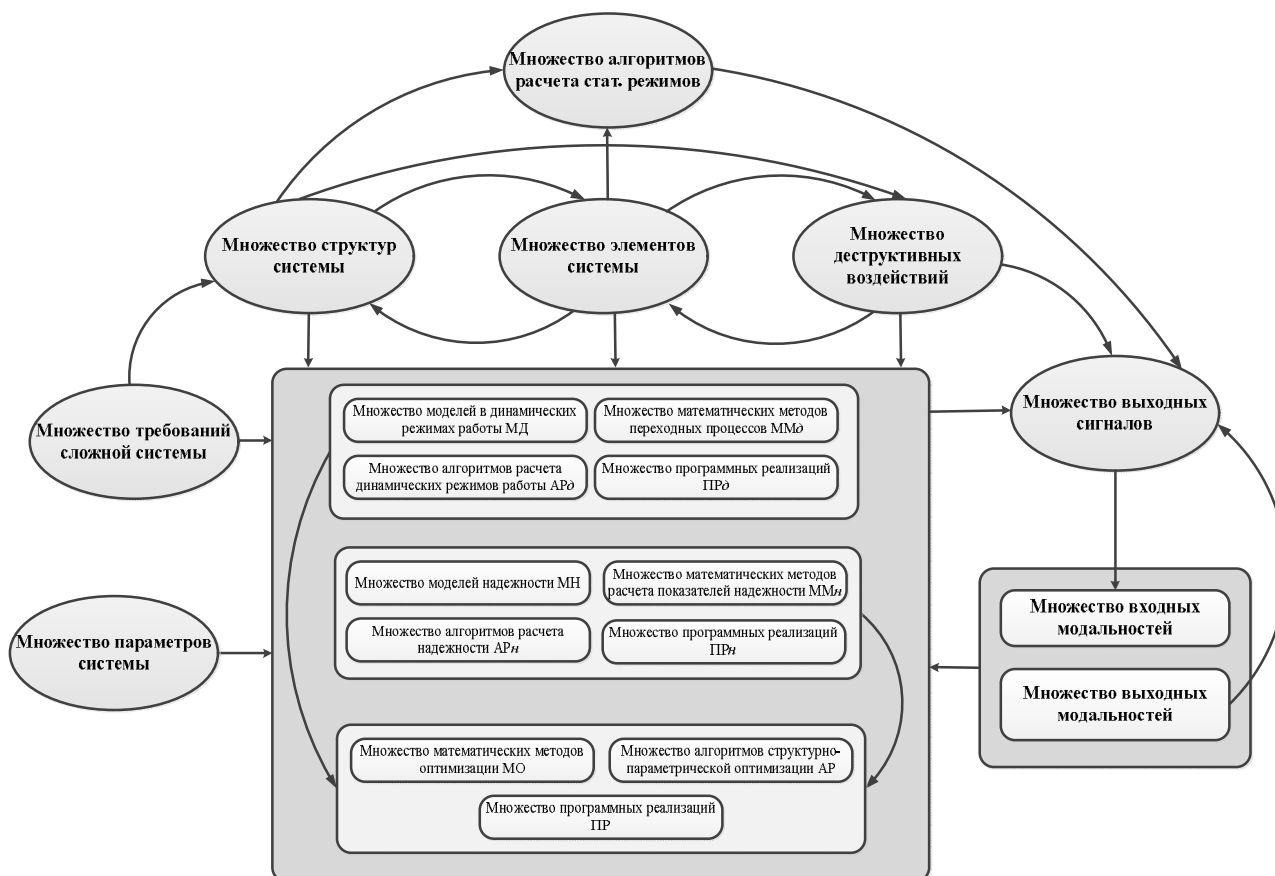


Рис. 1. Теоретико-множественная модель сложной нелинейной системы



Рис. 2. Обобщенная модель деструктивных воздействий на сложную нелинейную систему

Для установления и структурирования связей между множествами системы вводится множество $V = \{V_k\}$, состоящее из множества k связей между элементами различных уровней, представленных в виде морфологической таблицы (табл. 1) и сформированных на ее основе возможных путей (1).

Таблица 1. Морфологическая таблица множества связей между элементами различных уровней

Основное множество	Подмножества	Элементы		
A ₁ Множество структур	A ₁₁ Топологий соединения элементов	A ₁₁₁ Радиальная	A ₁₁₂ Магистральная	A ₁₁₃ Смешанная
	A ₁₂ Источников бесперебойного электропитания	A ₁₂₁ Буферная	A ₁₂₂ Буферная с конвертором	A ₁₂₃ Буферная с отделенной от нагрузки аккумуляторной батареей
	A ₁₃ Телекоммуникационного оборудования	A ₁₃₁ С расширением	A ₁₃₂ Без расширения	
A ₂ Множество элементов	A ₂₁ Преобразователи энергии	A ₂₁₁ Выпрямительные устройства	A ₂₁₂ Конверторы	
	A ₂₂ Источники энергии	A ₂₂₁ Аккумуляторные батареи	A ₂₂₂ Собственная электростанция	
	A ₂₃ Коммутация	A ₂₃₁ Автоматический ввод резерва	A ₂₃₂ Шкаф вводно-распределительный	A ₂₃₃ Рядовое распределительное устройство
	A ₂₄ Соединительные элементы	A ₂₄₁ Кабели	A ₂₄₂ Шины	
A ₃ Геометрическая топология пространства	A ₃₁ Габаритных размеров	A ₃₁₁ Компоненты	A ₃₁₂ Пространство	
	A ₃₂ Нормативных допусков			
A ₄ Защита	A ₄₁ Заземление			
	A ₄₂ Молниезащита			
	A ₄₃ Рядовая защита			

$$A_{11} \rightarrow A_{21} \rightarrow \dots \rightarrow A_{n2} \cup A_{11} \rightarrow A_{21} \rightarrow \dots \rightarrow A_{nk} \cup A_{11} \rightarrow A_{22} \rightarrow \dots \rightarrow A_{n1} \cup \dots \rightarrow A_{12} \rightarrow A_{22} \dots \rightarrow A_{n1} \cup A_{12} \rightarrow A_{2j} \rightarrow \dots \rightarrow A_{n2} \cup A_{1h} \rightarrow A_{22} \rightarrow A_{11} \rightarrow A_{n1} \dots \quad (1)$$

Для анализа функционирования системы при деструктивных воздействиях разработана обобщенная модель сложной нелинейной системы (рис. 3), позволяющая учесть уровни перенапряжений при различных деструктивных воздействиях с учетом разных вариантов соединения функциональных блоков (которые задаются сменой положений ключей S). Все состояния ключей описываются таблицей, пример которой представлен табл. 2 при деструктивном воздействии по входу.

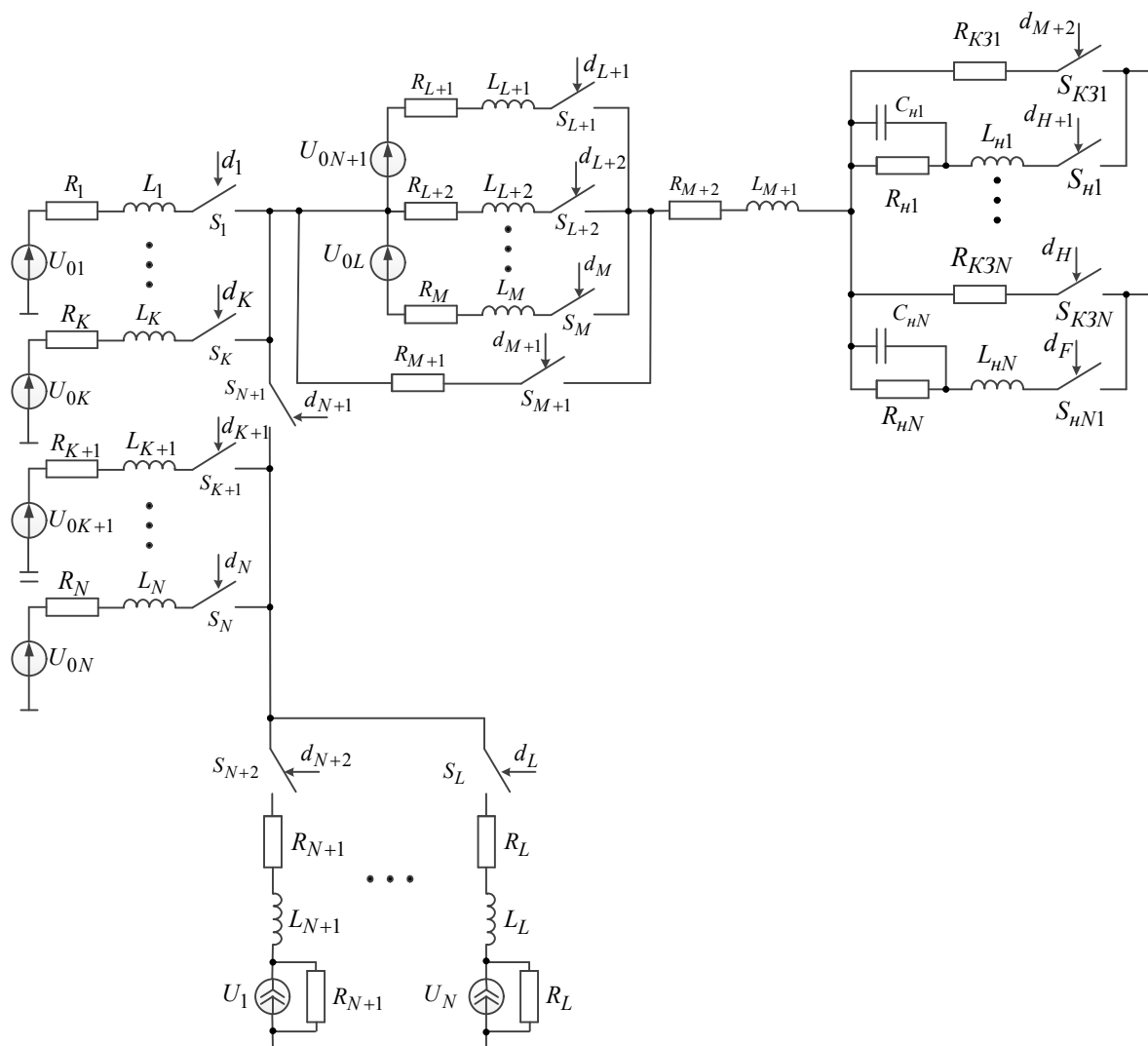


Рис. 3. Обобщенная модель для анализа деструктивных воздействий на сложную нелинейную систему

Таблица 2. Положения ключей при деструктивном воздействии по входу

Воздействие \ Режим	$d_1 \dots d_K$	$d_{K+1} \dots d_N$	d_{N+1}	$d_{N+2} \dots d_L$	$d_{L+1} \dots d_M$	d_{M+1}	$d_{M+2} \dots d_H$	$d_{H+1} \dots d_F$
	Состояния ключей							
Режим 1	0	0	1	$1 \rightarrow 0$	1	0	0	1
Режим 2	0	1	0	$0 \leftarrow 1$	1	0	0	1
Режим 3	0	0	1	1	1	0	0	1
...	...	...	...	...	...	...	...	...
Режим N	0	1	1	1	0	1	1	0

Результаты моделирования совокупности подмножеств моделей в динамических режимах работы (МД), математических методов при расчете динамического режима работы (ММ_д), алгоритмов расчета переходных процессов при деструктивных воздействиях $D_n(t_n)$ (AP_д) и программных реализаций при динамическом режиме работы (ПП_д) сводятся к множеству допустимых системотехнических решений Δ_α :

$$\Delta_\alpha = \left\{ \alpha = \langle m_d, ap_d, np_d, mm_d \rangle \mid m_d \in MD, mm_d \in MM_d, ap_d \in AP_d, np_d \in PP_d \right\}. \quad (2)$$

Для анализа надежности функционирования сложной нелинейной системы в аварийном и нормальном режимах работы приведена обобщенная схема (рис. 3), на основе различных конфигураций которой составляется оптимальная схема расчета надежности системы. Показатель $\lambda_{эл}$ является интенсивностью отказов элементов системы, а $\lambda_{сл}$ – интенсивность отказов соединительных линий системы. Общую схему надежности системы (рис. 3а) также можно представить в виде блока надежности, который, в свою очередь, можно соединять параллельным или последовательным способом (рис. 3б), укрупняя конечную схему.

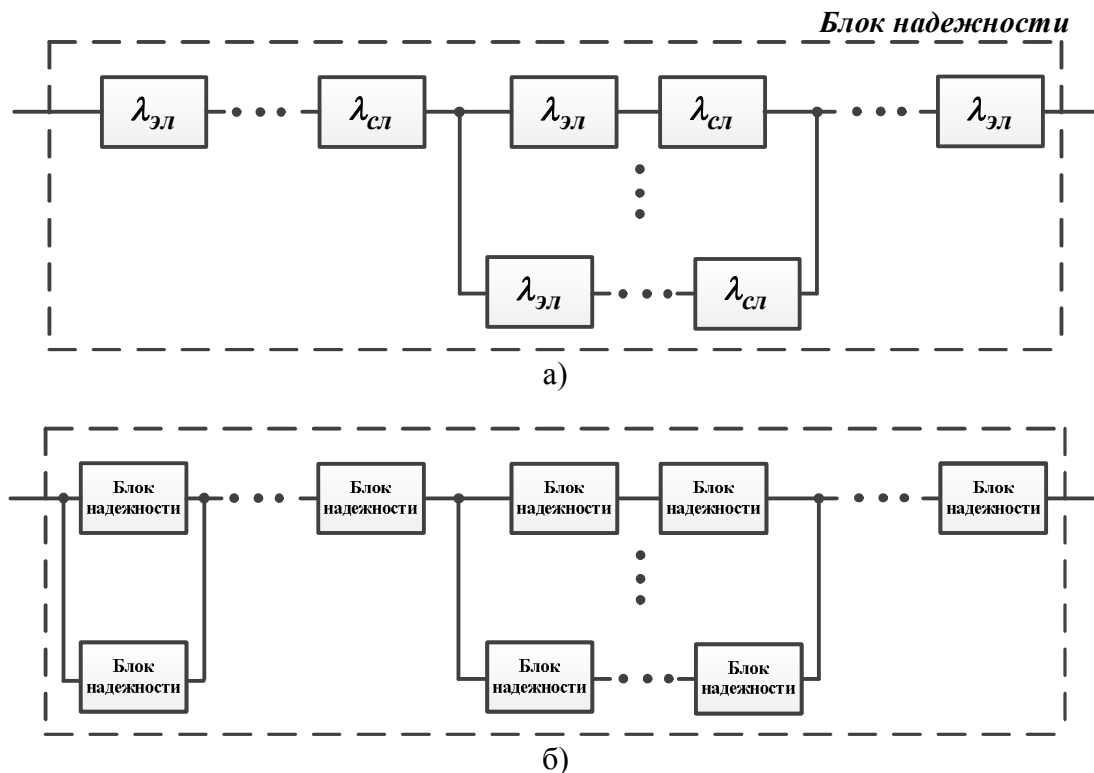


Рис. 3. Надежность системы

Результаты моделирования совокупности подмножеств моделей надежности (МН), математических методов расчета надежности (ММ_н), алгоритмов расчета параметров надежности (AP_н) и программных реализаций (ПП_н) сводятся к множеству допустимых системотехнических решений Δ_β :

$$\Delta_\beta = \left\{ \beta = \langle mn, ap_n, np_n, mm_n \rangle \mid mn \in MN, mm \in MM_n, ap \in AP_n, np \in PP_n \right\}. \quad (3)$$

Множество Δ_γ является множеством допустимых системотехнических решений, включающее подмножества математических методов оптимизации (МО), алгоритмов структурно-параметрической оптимизации (AP), программных реализаций (ПП):

$$\Delta_\gamma = \left\{ \gamma = \langle mo, np, ap \rangle \mid mo \in MO, ap \in AP, np \in PP \right\}. \quad (4)$$

3. Апробация

Предложенная теоретико-множественная модель была применена для сложной нелинейной системы, которой является опорная цифровая автоматическая телефонная станция (ЦАТС) МС240С с источником вторичного электропитания УЭП2-3. ЦАТС МС240С включает в себя 24 абонентских комплекта, центральный процессор, VoIP-шлюз, модуль ТМ.ІР, предназначенный для передачи голосовой и факсимильной информации. Потребляемая мощность центрального процессора составляет 15 Вт, VoIP-шлюза и модуля ТМ.ІР – 6 Вт и 6 Вт соответственно. УЭП2-3 питается от сети переменного тока 220 В и обеспечивает питание постоянным током ЦАТС до 48 А с выходным напряжением 43...57 В или 53...–69 В. К опорной ЦАТС МС240С подключаются блоки расширения (рис. 4), которые выполнены в виде каркаса станции МС240С с установленными модулями абонентских комплектов и модулем контроллера сопряжения. Коммутация соединений осуществляется ЦП.Е, установленным в управляющем каркасе.

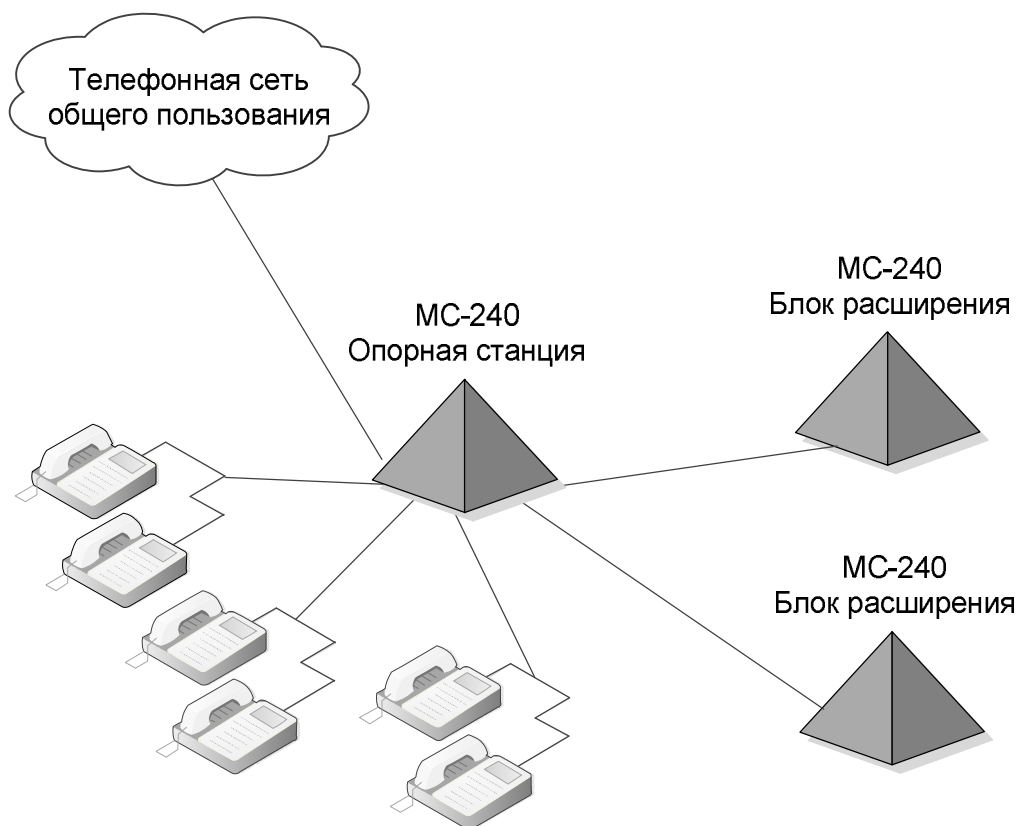


Рис. 4. Схема связи с расширением

Одним из основных методов оптимизации (МО на рис. 1) является структурно-параметрическая оптимизация, проводимая для отобранного оборудования по трем критериям: 1) техническим характеристикам; 2) условиям габаритных размеров; 3) показателям надежности и экономичности:

$$F_s = \min \left[\left(\min \sum_{k=1}^N F_{ОБ}^k \right) + \sum_{i=1}^M F_{ПР}^i \right]. \quad (5)$$

В результате проведения оптимизации были получены альтернативные решения выбора компонент системы и их размещения в пространстве, соответствующие выбранным критериям оптимизации с учетом нормативных требований к телекоммуникационной системе. Результаты лучших вариантов согласно трем этапам минимизации отображены на гистограммах

(рис. 5). В табл. 3 отображены некоторые численные характеристики лучших вариантов, где λ – показатель надежности, η – коэффициент полезного действия, C – суммарная стоимость оборудования, G – суммарный вес оборудования.

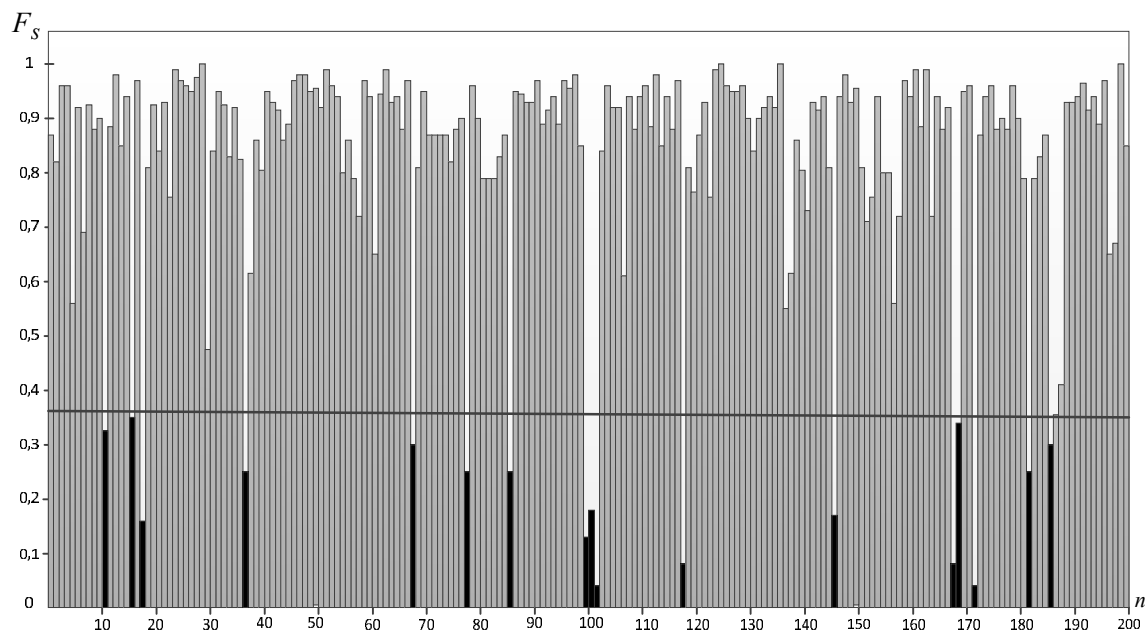


Рис. 5. Гистограммы минимизации оборудования

Таблица 3. Результаты лучших вариантов

№ варианта	11	15	17	36	68	78	99	100
λ	$8,5 \cdot 10^{-7}$	$9,1 \cdot 10^{-7}$	$8,3 \cdot 10^{-7}$	$8 \cdot 10^{-7}$	$8,02 \cdot 10^{-7}$	$8,10 \cdot 10^{-7}$	$7 \cdot 10^{-7}$	$8,5 \cdot 10^{-7}$
η	0,84	0,84	0,86	0,83	0,85	0,84	0,8	0,86
C	$1,880 \cdot 10^3$	$1,840 \cdot 10^3$	$1,11 \cdot 10^3$	$1,75 \cdot 10^3$	$1,811 \cdot 10^3$	$1,851 \cdot 10^3$	$1,74 \cdot 10^3$	$1,653 \cdot 10^3$
G	802	799	703	760	860	878	822	627
№ варианта	101	118	146	167	168	172	182	186
λ	$7,89 \cdot 10^{-7}$	$6,2 \cdot 10^{-7}$	$7,9 \cdot 10^{-7}$	$9,8 \cdot 10^{-7}$	$8,89 \cdot 10^{-7}$	$7,73 \cdot 10^{-7}$	$8,21 \cdot 10^{-7}$	$8,11 \cdot 10^{-7}$
η	0,84	0,85	0,84	0,83	0,85	0,87	0,84	0,82
C	$1,657 \cdot 10^3$	$1,855 \cdot 10^3$	$1,8 \cdot 10^3$	$1,940 \cdot 10^3$	$1,900 \cdot 10^3$	$1,572 \cdot 10^3$	$1,72 \cdot 10^3$	$1,763 \cdot 10^3$
G	758	850	832	826	760.4	658.6	880	780

При поиске глобального экстремума получена оптимальная структура ЦАТС (рис. 6).

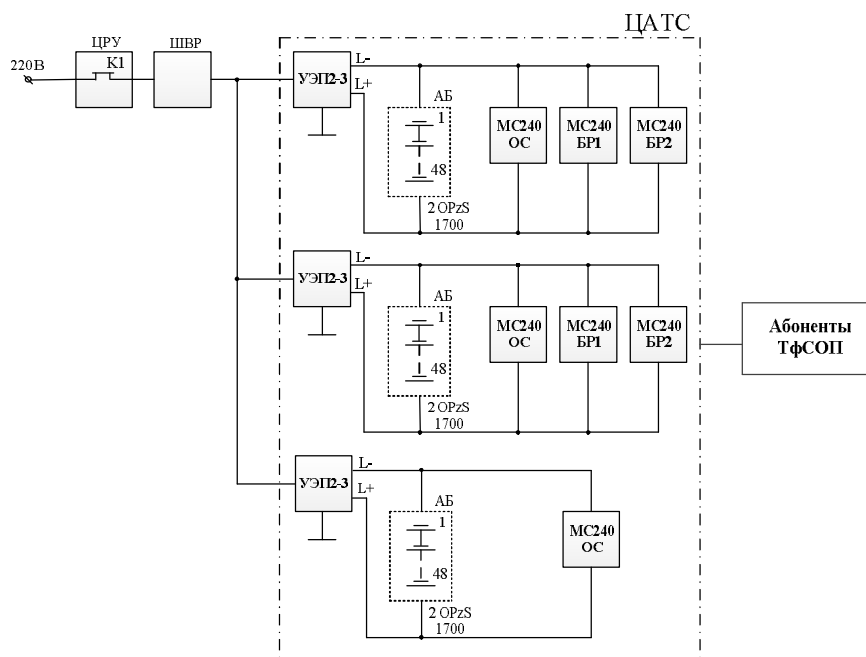


Рис. 6. Схема цифровой автоматической телефонной станции с блоками расширения

4. Выводы

Оптимальная структура цифровой автоматической телефонной станции включает в себя: аккумуляторные батареи 2ОПзС емкостью 1700А·ч в количестве 48 штук, выпрямительные устройства Eltek 3 стойки по 14 модулей, конвертор Eltek Micropack DC/DC 120. Суммарная надежность выбранного оборудования составляет $\lambda_{\text{сумм}} = 6.486 \cdot 10^{-6} \text{ 1/ч}$, коэффициент полезного действия $\eta = 0,85$, стоимость единицы массы оборудования $C = 1.855 \cdot 10^3$ условных единиц, суммарный вес оборудования $G = 850$ кг. По результатам проведенных расчетов цифровая автоматическая телефонная станция потребляет мощность 79.4419 кВА. Топология соединения элементов построена по радиально-магистральной схеме с выбором автоматов защиты DPX-630, DPX-250 и DPX-160, время срабатывания которых не превышает 0.9 с. Конфигураций сети – буферная система электропитания с конвертором.

Литература

1. Акимов С. В. Введение в морфологические методы исследования и моделирование знаний предметной области. Электронный ресурс. Режим доступа: <http://structuralist.narod.ru/dictionary/morphbox.htm>.
2. Акимов С. В. Structuralist – язык моделирования морфологического множества // 56-я НТК СПбГУТ. СПб.: СПбГУТ, 2004. С. 75.
3. Волкова В. Н. Денисов А. А. Теория систем и системный анализ, 2-е изд., перераб. и доп. М.: Издательство Юрайт, 2015. 616 с.
4. Одрин В. М., Картавов С. С. Морфологический анализ систем. Построение морфологических таблиц. К.: Наукова думка, 1977. 148 с.
5. Антонов А. В. Системный анализ. М.: Высшая школа, 2004. 454 с.
6. Губанов В. А., Захаров В. В., Коваленко А. Н. Введение в системный анализ. Л.: ЛГУ, 1988. 228 с.

Статья поступила в редакцию 12.04.2019.

Губкина Валерия Руслановна

старший преподаватель кафедры САПР СибГУТИ (630102, Новосибирск, ул. Кирова, 86),
тел. (383) 269-82-59, e-mail: vg@sibsutis.ru.

**Set-theoretic model for solving the problem of system analysis and synthesis
in telecommunications**

V. Gubkina

A set-theoretic model for solving the problem of system analysis and synthesis in a complex non-linear system, taking into account the input parameters, the requirements for the output characteristics, the variability of the architecture of the system, external destructive effects on it is considered. The results of testing approbated on digital automatic telephone exchange are presented.

Keywords: complex nonlinear dynamic system, system analysis and synthesis, set-theoretic model, electromagnetic compatibility, digital automatic telephone exchange.

Режимы функционирования многофункциональной бортовой РЛС БЛА малой и средней дальности

Е. М. Ильин, Д. А. Репников, В. Ю. Савостьянов, О. Ф. Самарин,
А. И. Полубехин, А. Г. Черевко

Рассмотрены основные режимы функционирования многофункциональной бортовой радиолокационной системы беспилотного летательного аппарата (МБРЛС БЛА) малой и средней дальности. Эффективное использование оператором БЛА реализованных в МБРЛС возможностей потребовало разработки соответствующей тактики применения тех или иных режимов при выполнении конкретных разведывательно-ударных задач (обнаружения и идентификации, мониторинга, сопровождения) по различным объектам. Представлены некоторые возможные сценарии решения этих задач, а также результаты натурных испытаний МБРЛС

Ключевые слова: многофункциональная бортовая радиолокационная станция, радиолокационная целевая нагрузка, беспилотный летательный аппарат, режимы функционирования, тактика, разведывательно-ударные задачи, картографирование, селекция целей.

1. Введение

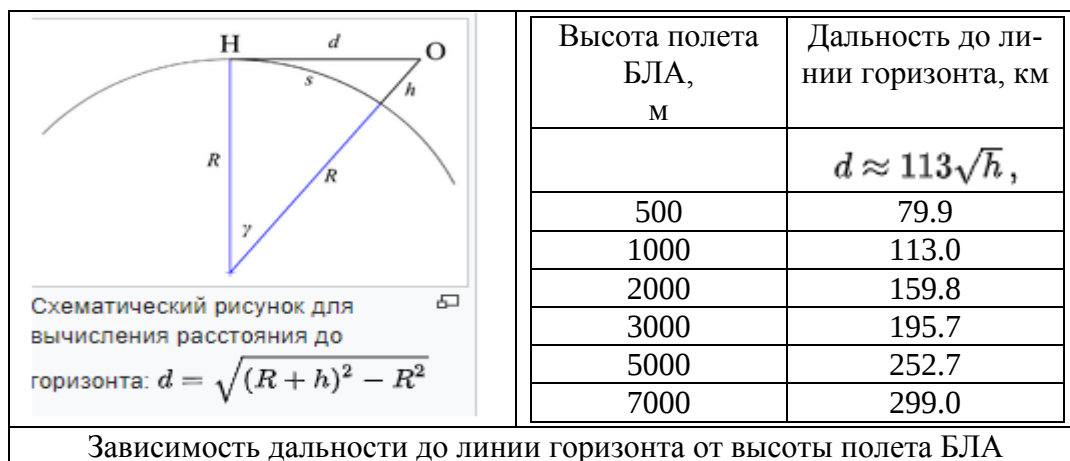
Комплексы с беспилотными летательными аппаратами (КБЛА) находят самое широкое применение при решении народнохозяйственных и военных задач.

На разных стадиях разработки находятся отечественные БЛА малой и средней дальности (МД и СД) типа «Корсар», «Иноходец», «Форпост-Р» и др. (табл. 1). Интенсивно идет разработка радиоэлектронного оборудования для этих БЛА, в частности, оптико-электронных и радиолокационных средств, средств радиотехнической разведки. При этом допустимая масса создаваемых радиоэлектронных средств разведки, обеспечивающая их применение в составе перспективных БЛА, не может превышать 30–60 кг, поскольку масса радиоэлектронных средств, как правило, не превышает 30 % от общей массы целевой нагрузки.

Таблица 1 Технические характеристики отечественных БЛА МД и СД

Наименование БЛА	Высота полета, км	Скорость, км/ч	Дальность полета, км	Масса целевой нагрузки, кг
БЛА «Корсар»	до 7	80...220	250	до 45
БЛА «Иноходец»	до 7.5	120...150	250	до 55
БЛА «Форпост»	до 5.8	200	250	до 50
БЛА «Платформа»	до 3	70	350	до 50
БЛА «Катран»	до 4	130	250	до 120

Сформулируем общие требования к многофункциональной радиолокационной целевой нагрузке (МРЛЦН) КБЛА МД и СД.



При предельной высоте полета БЛА 7 км дальность до линии горизонта (ДЛГ) составляет около 300 км. При средней высоте полета БЛА 3...5 км ДЛГ лежит в пределах 190...250 км. На суше ДЛГ еще зависит от рельефа местности. Эти параметры определяют основное требование к МРЛЦН по дальности лоцирования наземных и морских объектов. В режимах картографирования дальность определяется требуемым разрешением и для приемлемых разрешений лежит в пределах до 100 км.

Исходя из вышеназванных задач и требований к КБЛА формируются требования по составу и тактико-техническим характеристикам (ТТХ) целевых нагрузок КБЛА и, в частности, радиолокационных целевых нагрузок:

- функционирование в различных частотных диапазонах в пределах 0.1...100 ГГц;
- картографирование/мониторинг земной/морской поверхности на дальностях до 100 км с разрешением от 0.25 м;
- мониторинг окружающего пространства с полосой обзора до 60 км;
- обнаружение и измерение координат неподвижных радиоконтрастных и движущихся наземных (надводных) объектов на дальностях до 300 км;
- радиомониторинг окружающего пространства в рабочей полосе частот МРЛЦН;
- оценка метеобстановки с дальностью обнаружения метеопреобразований с отражаемостью 40 дБЗ до 200 км.

В результате исследований, проведенных МГТУ имени Н. Э. Баумана совместно с соисполнителями, разработана многофункциональная малогабаритная бортовая РЛС (МБРЛС) Ku-диапазона (рис. 1), предназначенная для применения в составе разведывательно-ударных комплексов с беспилотными летательными аппаратами оперативно-тактического звена малой и средней дальности [1–3].

Выбор Ku-диапазона длин волн обусловлен тем, что позволяет получать детальные радиолокационные изображения объектов в любое время суток, в любых погодных условиях, на поле боя в условиях дымовых и пылевых завес.

Прототип этой МБРЛС [4] демонстрировался на III-й Научно-практической конференции «Роботизация Вооружённых Сил Российской Федерации» и Международном военно-техническом форуме «Армия-2018».

Созданная МБРЛС отличается от прототипа:

- более эффективным функциональным программным обеспечением (ФПО) за счёт введения новых функций в режимах работы;
- безредукторным приводом антенной системы, обеспечивающим углы прокачки антенны до $\pm 140^\circ$ по азимуту и до $-45...+5^\circ$ по наклону, а также погрешность позиционирования луча антенны не более 2 угл. мин.;
- увеличенными значениями частоты дискретизации аналого-цифрового преобразователя (с 160 до 320 МГц) и полосы пропускания цифрового приёмника (с 20 до 40 МГц);
- повышенными значениями производительности центрального процессора (с 16 до 38.4 Гфлопс), объёма оперативной памяти (с 4 до 16 Гбайт) и объёма накопителя SSD БЦВМ

(с 32 до 128 Гбайт);

- аппаратно-программной интеграцией подсистемы микронавигации (ПМН) в состав РЛС;
- сниженной массой РЛС (с 38.5 до 30.5 кг без ПМН).

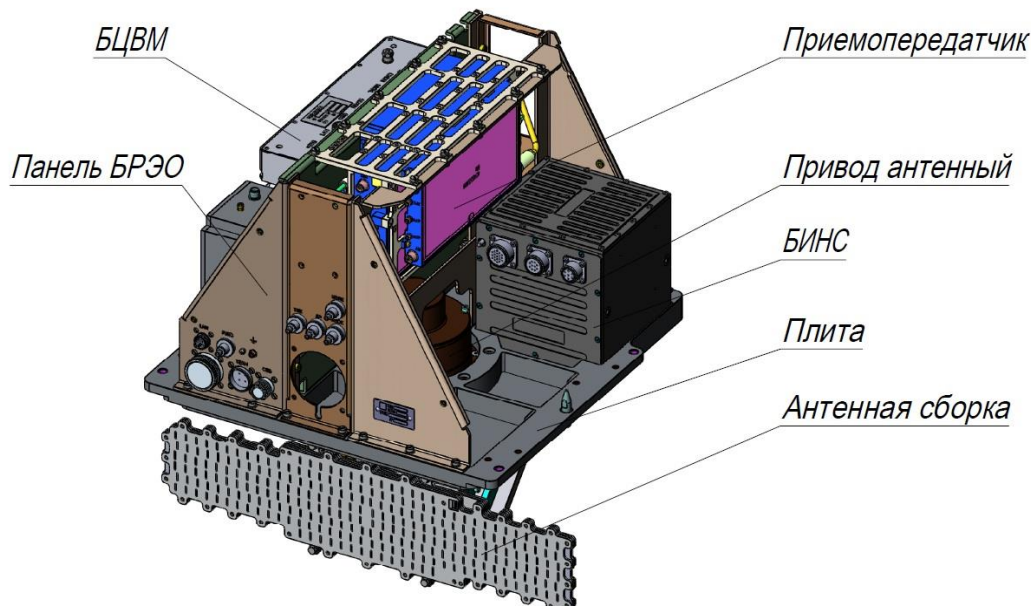


Рис. 1. МБРЛС Ку-диапазона длин волн

2. Основные режимы функционирования МБРЛС БЛА МД и СД

Основные режимы работы МБРЛС:

- картографирование с реальным лучом или с синтезированием апертуры антенны при полосовом, секторном или телескопическом обзоре с формированием радиолокационных изображений (РЛИ) подстилающей поверхности размером 1024×1024 пикселей с возможностью оперативного изменения разрешения, дальности и угла наблюдения по азимуту;
- селекция наземных (надводных) движущихся целей (СНДЦ), в том числе с одновременным формированием РЛИ подстилающей поверхности;
- оценка метеообстановки с формированием РЛИ сечений метеообразований (вертикальных, горизонтальных) размером 512×512 пикселей с обнаружением и индикацией зон опасной турбулентности и опасного низковысотного «сдвига ветра»;
- измерение наклонной дальности до поверхности по заданному угловому целеуказанию;
- информационное обеспечение полёта на малых высотах с формированием РЛИ сечений рельефа впередилежащей местности (вертикальных, горизонтальных, плановых) размером 512×512 пикселей;
- радиомониторинг в пределах полосы рабочих частот МБРЛС с формированием РЛИ азимутально-спектральных карт внешних излучений размером 512×512 пикселей с возможностью оперативного изменения оператором угла наблюдения по наклону;
- встроенная система контроля с возможностью выполнения текущих, расширенных или диагностических проверок.

В соответствии с требованиями к оперативному обнаружению и сопровождению объектов разработанные аппаратная и программная платформы МБРЛС позволяют формировать целевую информацию (радиолокационные изображения подстилающей поверхности, сечений метеообразований и рельефа местности, координаты и параметры движения объектов) непосредственно в МБРЛС в реальном масштабе времени без вмешательства оператора

наземного пункта управления (НПУ). При этом в процессе формирования сообщения в его заголовок записывается информация (геодезические координаты центра РЛИ, азимут, межпиксельное расстояние), позволяющая автоматизировать «привязку» целевой информации к топографической карте. Сформированная целевая информация с борта БЛА по радиоканалу передаётся на НПУ, где осуществляется её визуализация, дешифрирование и «сколка» объектов.

Для эффективного решения различных тактических задач в МБРЛС используется секторный, полосовой, телескопический или многократный телескопический обзор поверхности, любой из которых, в принципе, позволяет производить просмотр больших участков местности, представляющих интерес (рис. 2).

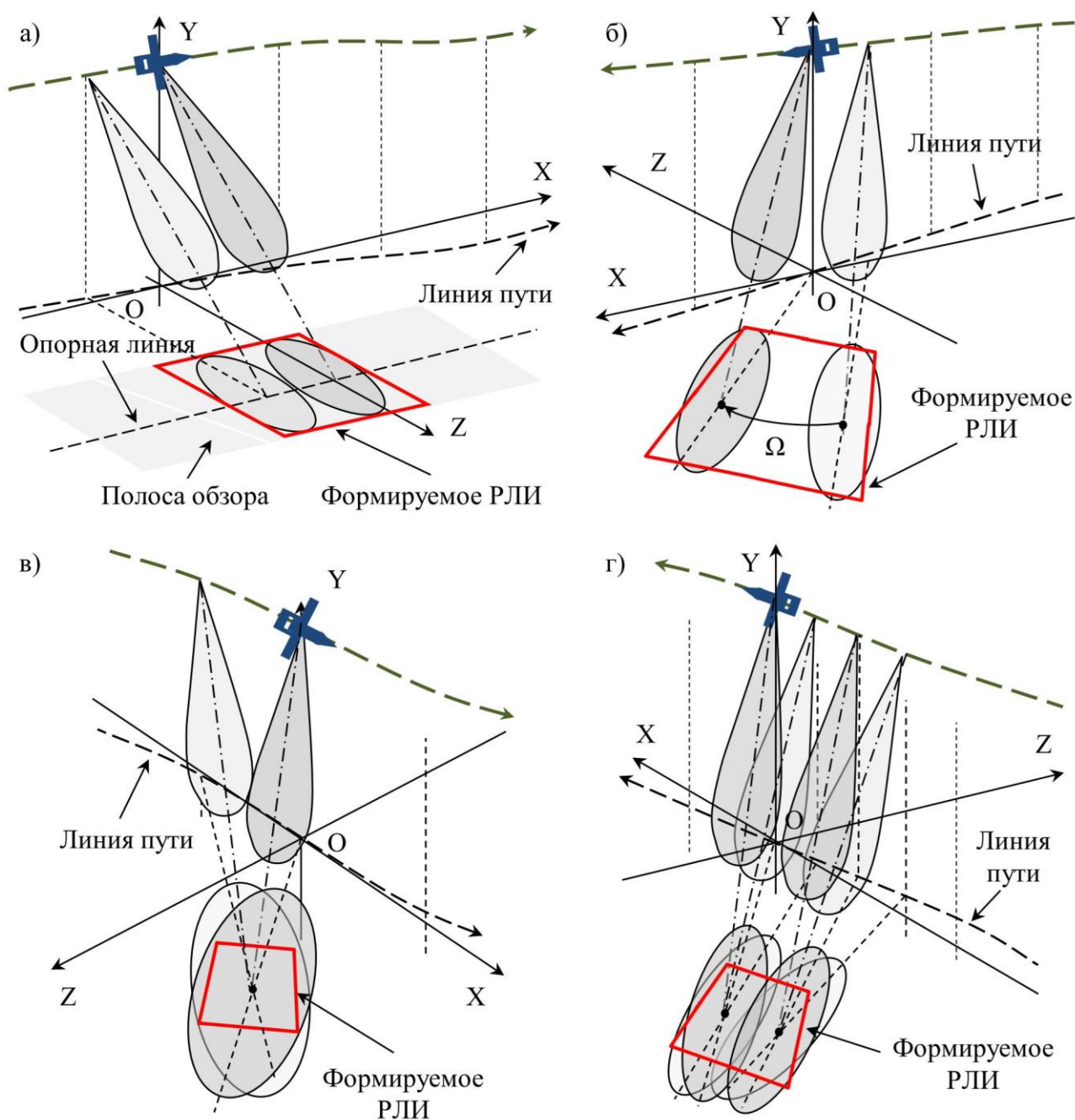


Рис. 2. Виды обзора поверхности

В случае полосового обзора БЛА перемещается вдоль некоторой линии по заданному курсу, при этом антенна МБРЛС стабилизирована в пространстве, и в течение длительного времени ведётся обзор местности с непрерывным формированием РЛИ в виде ленты (рис. 2а).

Секторный обзор используется для расширения по азимуту участка местности, помещающегося в кадр РЛИ. Антенна РСА сканирует в горизонтальной плоскости с некоторой угловой скоростью Ω , которая определяется требуемой разрешающей способностью, путевой скоростью носителя и углами наблюдения. Траектория центральной точки на поверхности стремится к окружности, при этом форма зоны обзора оказывается близка к форме сектора (рис. 2б).

Телескопический (прожекторный) обзор характеризуется тем, что в течение всего времени наблюдения обеспечивается наблюдение участка местности в окрестности некоторой точки (рис. 2в). Этот вид обзора позволяет реализовать наиболее высокое разрешение. Недостатком телескопического обзора является ограниченность размера картографируемого участка поверхности.

При многократном телескопическом обзоре РЛИ формируется из нескольких фрагментов – парциальных кадров (ПК), в каждом из которых осуществляется телескопический обзор (рис. 2г). Размер картографируемого участка поверхности здесь может быть произвольным, способы стыковки ПК в изображении также могут быть любыми. Алгоритм обработки сигналов при этом будет один и тот же, а РЛИ заданного разрешения может быть получено за минимальное физически реализуемое время. Становится возможным оперативно изменять разрешение в РЛИ, дальность наблюдения и углы картографирования, решая различные тактические задачи. Именно этот вид обзора, часто используемый в РЛС фронтовой авиации, представляется наиболее удачным вариантом с точки зрения применения в МБРЛС для БЛА малой и средней дальности.

В МБРЛС используются частотно-манипулированные зондирующие сигналы с быстрой перестройкой несущей и псевдослучайным изменением закона следования частот, а также внутриимпульсная линейно-частотная модуляция с изменяемыми значениями девиации частоты, длительности импульсов и количества частот. Применение таких сигналов позволило не только снизить требования к аппаратуре, но и повысить скрытность и помехозащищённость радиолокатора. Получение требуемого высокого разрешения по дальности (до 0.25 м) достигается путём синтеза спектра частот.

Для обеспечения функционирования режимов в состав МБРЛС введена подсистема микронавигации (ПМН), созданная на базе БИНС ГЛ-100 и установленная на общем основании с МБРЛС. ПМН комплексирована с приёмником глобальной навигационной спутниковой системы, а также навигационной системой БЛА и МБРЛС. По данным, получаемым от ПМН, в реальном масштабе времени в ФПО МБРЛС осуществляется управление лучом антенны, периодом повторения зондирующих импульсов, положением зоны приёма и параметрами опорных функций.

3. Результаты натурных испытаний МБРЛС БЛА МД и СД

С экспериментальным образцом в 2017 – 2018 гг. было проведено более 20 натурных испытаний в составе летающей лаборатории, в результате которых на борту в реальном масштабе времени было получено более тысячи РЛИ подстилающей поверхности в различных режимах работы, при различных разрешениях, дальностях, углах и видах обзора. Использование высокоточной навигационной информации позволило обеспечить стыковку фрагментов РЛИ и их фокусировку, а также точное определение координат объектов.

В качестве примера на рис. 3–7 приведены РЛИ, полученные на борту летающей лаборатории в реальном масштабе времени при различных видах обзора с линейным разрешением от 0.25 до 3.8 м, в «привязке» к спутниковым оптическим фотоснимкам.

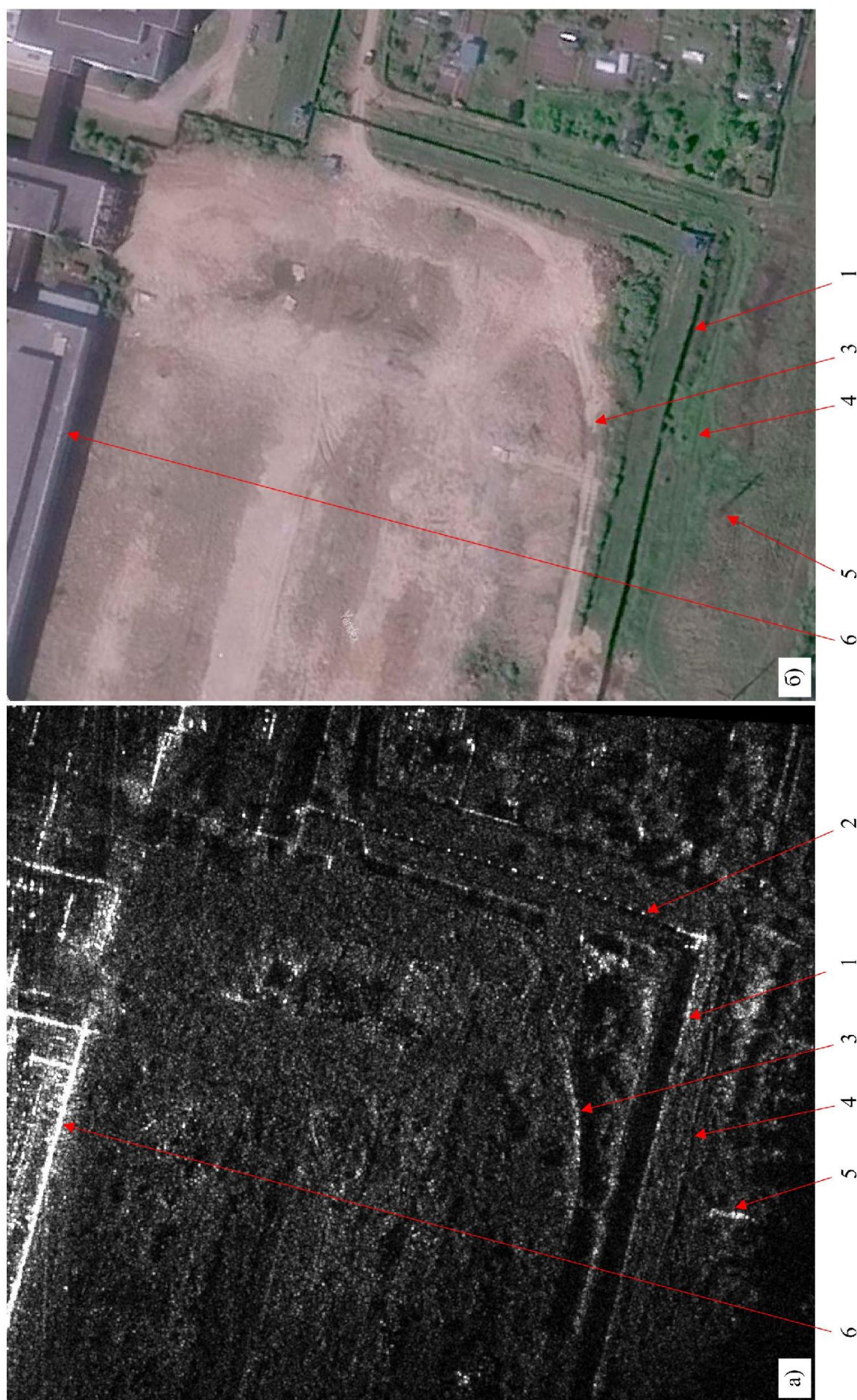


Рис. 3. РЛИ с разрешением 0.25 м (а) и спутниковый фотоснимок (б) (рисунок, выполненный в цвете, можно посмотреть на сайте журнала)



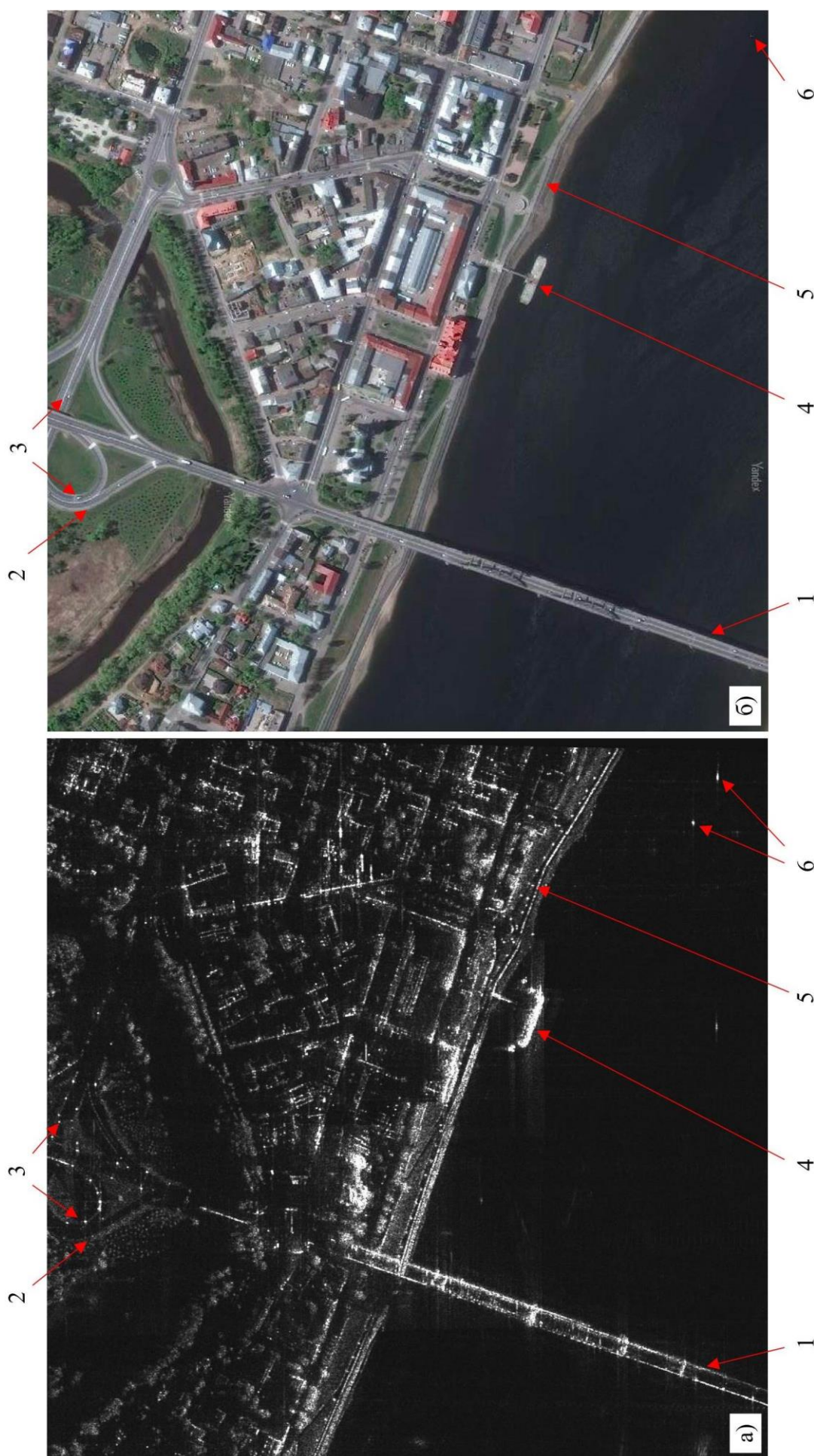


Рис. 5. РЛИ с разрешением 1 м (а) и спутниковый фотоснимок (б) района Волжского моста (г. Рыбинск).

Режим КРТ-СА. Полосовой обзор. Расстояние 6.5 км. Высота 1000 м.

1 – автомобильный мост, 2 – автомобильная развязка, 3 – автомобиль, 4 – причал, 5 – столбы освещения, 6 – бакен (рисунок, выполненный в цвете, можно посмотреть на сайте журнала)

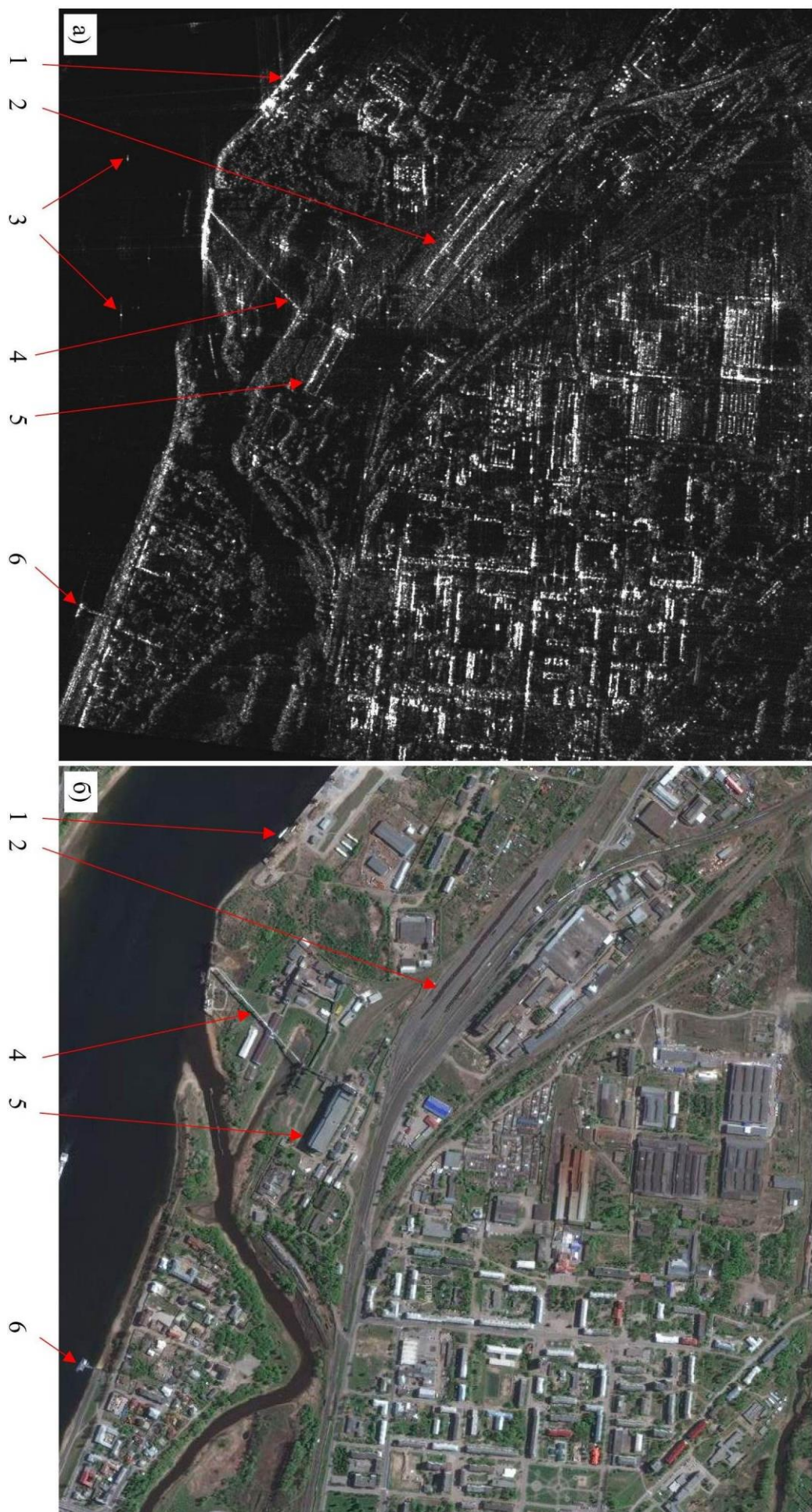


Рис. 6. РЛИ с разрешением 1,9 м (а) и спутниковый фотоснимок (б) района Зачеремульский (г. Рыбинск). Режим КРТ-СА. Секторный обзор. Расстояние 7,5 км. Высота 1200 м.
1 – причал грузового порта, 2 – железнодорожные составы, 3 – баки, 4 – транспортёр зерна, 5 – элеватор, 6 – причал (рисунок, выполненный в цвете, можно посмотреть на сайте журнала)

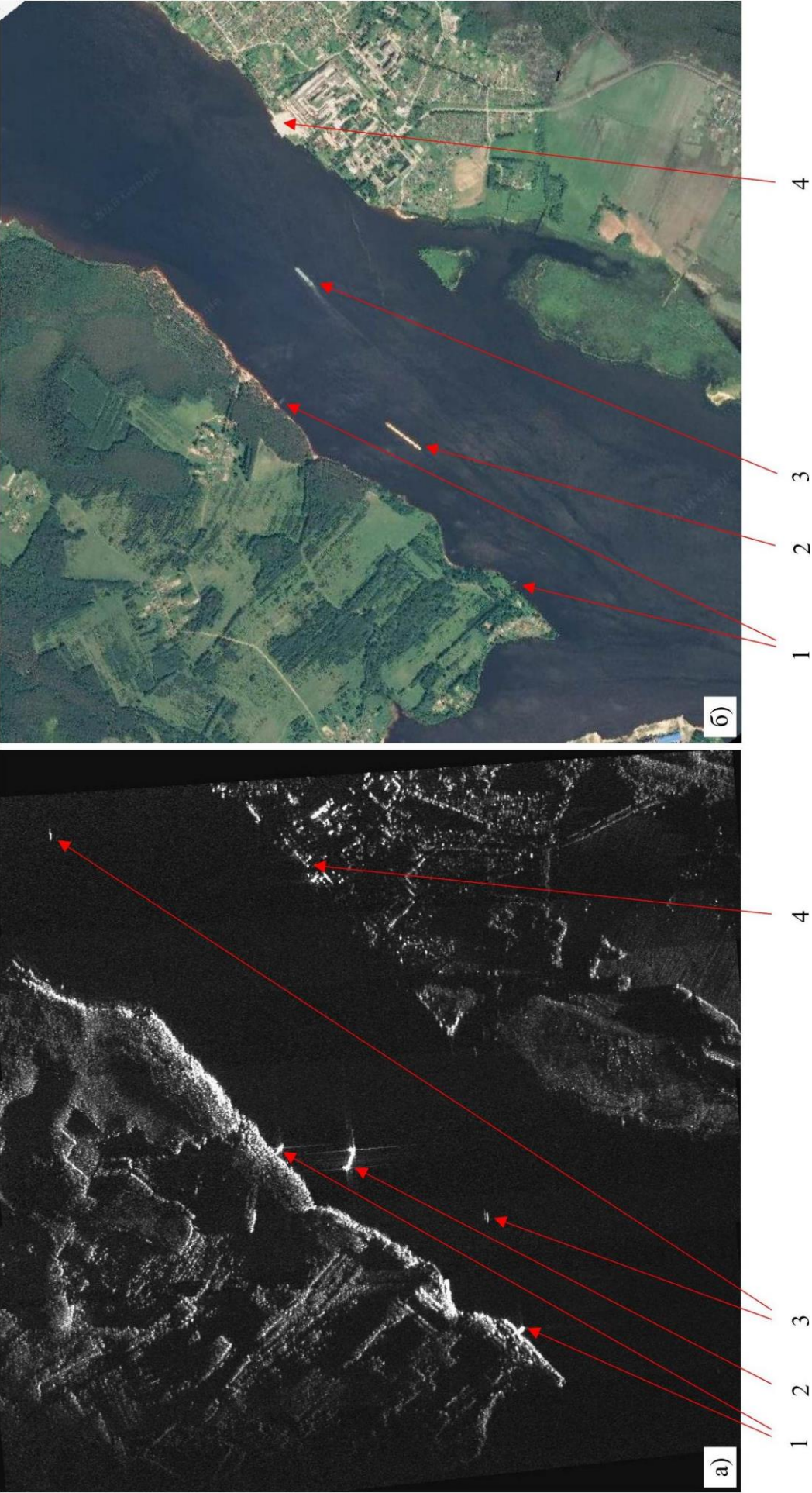


Рис. 7. РЛИ с разрешением 3.8 м (а) и спутниковый фотоснимок (б) района о. Юршинский (г. Рыбинск).
Полосовой обзор со сверхдлинным накоплением. Расстояние 25.3 км. Высота 1300 м.
1 – пристань, 2 – баржа, 3 – катер, 4 – причал грузового порта
(рисунок, выполненный в цвете, можно посмотреть на сайте журнала)

На рис. 8 показано результирующее РЛИ с разрешением 1.9 м, собранное из нескольких полученных на борту изображений, также в «привязке» к спутниковому оптическому фотоснимку района.

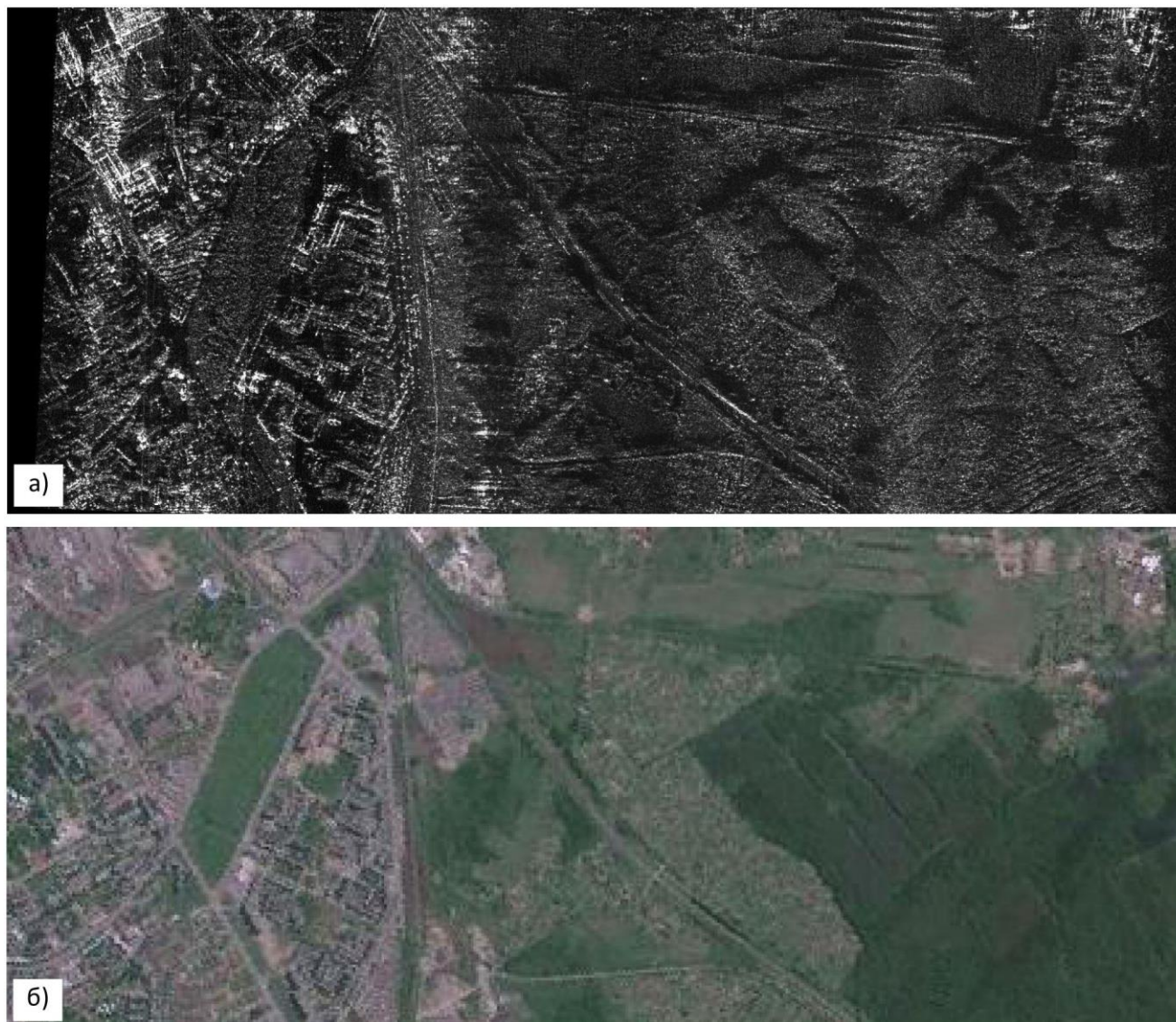


Рис. 8. Результирующее РЛИ с разрешением 1.9 м (а) и спутниковый фотоснимок (б) района Заволжье-2 (г. Рыбинск). Полосовой обзор. Расстояние 10.7 км. Высота 2400 м (рисунок, выполненный в цвете, можно посмотреть на сайте журнала)

МБРЛС Ку-диапазона не имеет аналогов в России. Известные подобные системы зарубежного производства не обладают таким количеством реализуемых функций. Большие информационные возможности и потенциал разработанной МБРЛС обеспечивают решение широкого круга задач для БЛА малой и средней дальности.

4. Тактика решения разведывательно-ударных задач БЛА МД и СД

Для эффективного использования реализованных в МБРЛС возможностей оператором НПУ необходима разработка соответствующей тактики применения тех или иных режимов при выполнении конкретных разведывательно-ударных задач (обнаружения и идентификации, мониторинга, сопровождения) по различным объектам. Далее представлены некоторые возможные сценарии решения этих задач.

Задача обнаружения радиоконтрастных объектов на поверхности в общем случае решается с помощью режимов картографирования и СНДЦ с различными разрешениями и видами

обзора. Однако тактика применения этих режимов в МБРЛС зависит от количества, размеров и мобильности целей. Рассмотрим несколько возможных вариантов (рис. 9–12).

Цель наземная множественная или одиночная, крупно- или среднеразмерная – аэродром, площадка, используемая для взлёта и посадки самолётов (вертолёт), мост, железнодорожный состав, здание, РЛС, позиция пусковых установок ракет, огневая позиция артиллерии, сосредоточение войск, боевой техники, материальных средств, пунктов погрузки и выгрузки, колонна техники, танк, тягач и т.п.

Если местоположение данного объекта неизвестно, то его поиск целесообразно начинать с применения режимов картографирования с секторным обзором и разрешением 7.5...30 м без стабилизации центра зоны. Так как размеры наблюдаемого участка поверхности при этом составляют от 7.5×7.5 км до 30×30 км, то, перемещая зону обзора путём изменения расстояния и азимутального угла до её центра, можно достаточно быстро осуществить как ориентирование на местности, так и собственно обнаружение объекта. Если местоположение объекта примерно известно, то достаточно будет одного РЛИ. Схематично эта процедура изображена на рис. 9.

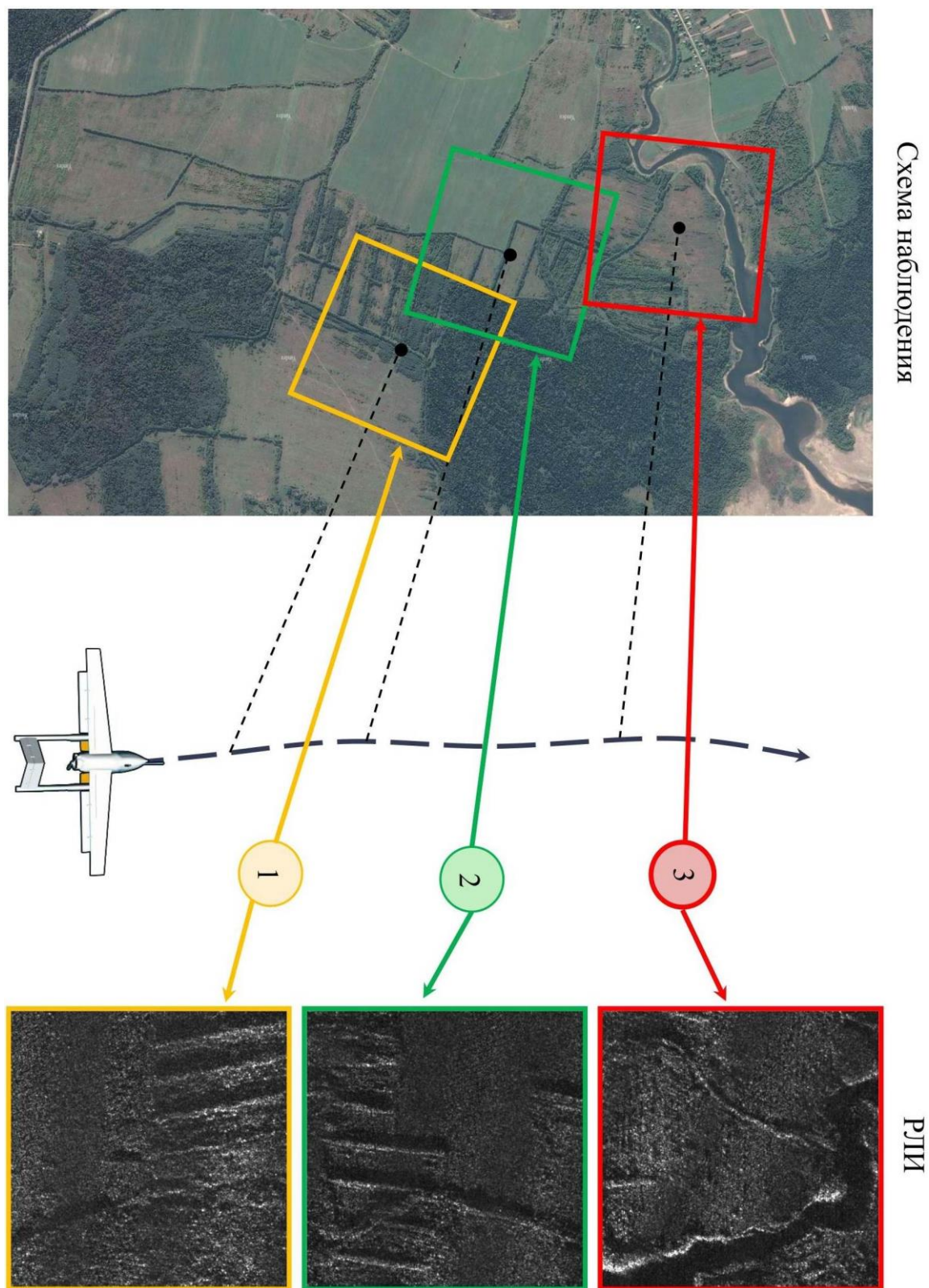
Дальнейшие действия оператора – это идентификация объекта, выполняемая в тех же режимах с секторным обзором, но со стабилизацией центра зоны и с повышением разрешения до 1...1.9 м. Схематично эта процедура изображена на рис. 10.

Если цель неподвижная, то её обнаружение может производиться только по радиолокационной контрастности (соотношению ЭПР цели и фона) в РЛИ. Если же цель мобильная, то появляется возможность её поиска ещё и по наличию радиальной скорости движения, что значительно повышает вероятность обнаружения. Тактика поиска целей будет такой же, но одновременно с формированием РЛИ можно будет производить СНДЦ. При этом оператор на экране индикатора сможет наблюдать отметки от движущихся целей на фоне обычного РЛИ.

Цель наземная множественная или одиночная, малоразмерная – артиллерийское орудие, пусковая установка ракет, автомобиль, мотоцикл, колонна бойцов, одиночный боец и т.п. Как правило, местоположение такого объекта примерно известно заранее, иначе его найти оперативно будет достаточно трудно. В остальном тактика поиска будет аналогична рассмотренной ранее (включая селекцию движущихся целей), но с более высоким разрешением – 1.9...7.5 м при поиске и 0.25...1 м при идентификации. В случае использования самого высокого разрешения вид обзора может измениться с секторного на телескопический (прожекторный). В МБРЛС такой переход осуществляется автоматически в зависимости от установленных значений разрешения и дальности.

Цель надводная множественная или одиночная, крупно- или среднеразмерная – корабль (или группа кораблей), катер. Тактика поиска такой цели во многом аналогична наземной, но с более грубым разрешением – 15...60 м при поиске и 3.8...15 м при идентификации. Отметим также, что ввиду низких значений ЭПР и её одинаковости для водной поверхности в пределах РЛИ можно выполнять картографирование не только с синтезированием апертуры антенны, но и с реальным лучом (в том числе с обнаружением луча по информации от разностно-азимутального канала). Разрешение по азимуту при этом ухудшится, но зато возрастёт скорость обзора. Если цель движется, то она эффективно будет обнаруживаться в режиме СНДЦ.

Цель надводная множественная или одиночная, малоразмерная – шлюпка, перископ подводной лодки, плывущий человек. В этой ситуации обнаружение объекта может осуществляться как с использованием режимов картографирования (если цель неподвижна) или СНДЦ (если цель движется) при разрешении 0.25...3.8 м, так и по скоростному портрету морской поверхности, который позволяет визуально наблюдать её аномалии с точностью до единиц сантиметров в секунду. Однако после обнаружения объекта его идентификацию можно выполнить только при детальном разрешении. Так как поиск подобных объектов, как правило, осуществляется в конкретном районе (например, вдоль береговой полосы), то здесь можно применять как секторный, так и полосовой обзор.



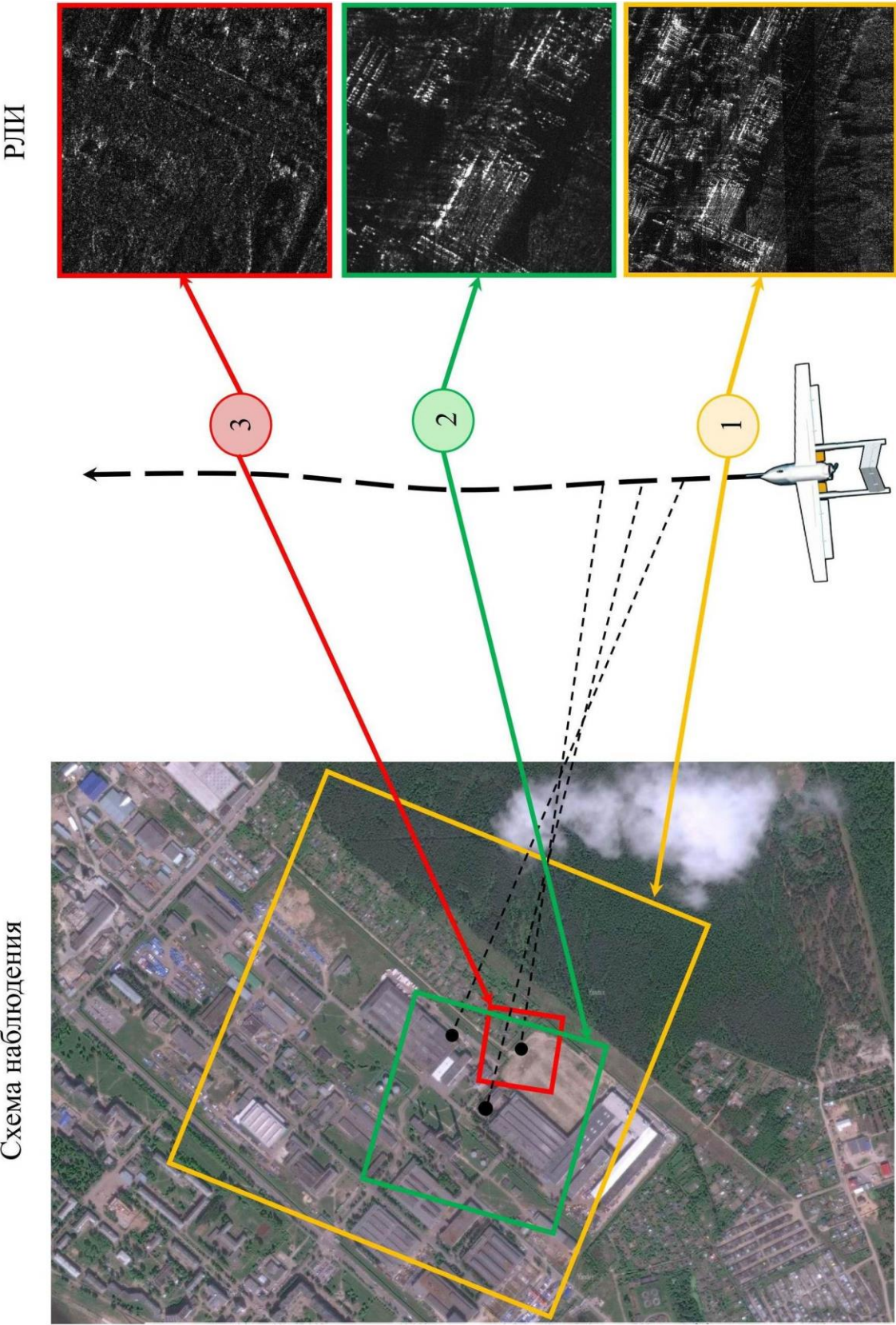


Рис. 10. Формирование РЛИ при идентификации объекта
(рисунок, выполненный в цвете, можно посмотреть на сайте журнала)

Задача радиолокационного мониторинга прибрежной полосы, городской, сельской местности или труднодоступных районов успешно решается в режимах картографирования с полосовым обзором (КРТ-ПО). Схематично эта процедура изображена на рис. 11. Поиск цели (или группы целей) целесообразно начинать с разрешений 1.9...15 м при углах картографирования 30...60° относительно БЛА по левому или правому борту. При появлении интересующих объектов оператор может перейти к детальным разрешениям (0.25...1 м) при углах картографирования 60...90°. Такая тактика позволит оперативно (в реальном времени за один пролёт БЛА) произвести не только обнаружение, но и идентификацию цели. Особенностью режимов КРТ-ПО является тот факт, что для формирования фрагментов РЛИ необходимо каждый раз перемещаться на определённое расстояние, причём для разрешений 3.8 м и более этот интервал оказывается намного длиннее, чем требуется для обеспечения заданного разрешения. В этом случае время облучения объекта может быть значительно больше, чем при секторном обзоре, а следовательно, может быть существенно увеличена и дальность картографирования. В МБРЛС такое сверхдлинное накопление сигнала реализовано и при необходимости может быть использовано оператором.

Задача распознавания объекта совершенно необходима для работы разведывательно-ударного комплекса с БЛА после обнаружения цели. И хотя распознавание при выполнении задания может использоваться оператором всего один раз, его роль в принятии решения очевидна. Решение задачи распознавания требует от МБРЛС обеспечения такой детализации РЛИ, при которой на изображении укладывается не менее чем 10...15 элементов разрешения по линейному размеру объекта или 100...200 элементов на площади изображения типовых сложных (групповых) целей. Известно, что для распознавания всех типов наземных военных объектов по РЛИ требуется потенциальное пространственное разрешение РЛС не хуже 0.3 м. Однако осуществление задачи распознавания требует не только получения детальных РЛИ, но также наличия достаточного количества эталонных изображений, сформированных при различных ракурсах и разрешениях, для всех интересующих оператора объектов. Следовательно, необходимы достаточно объёмная база данных для хранения этих эталонов и мощный вычислитель для эффективной работы алгоритмов распознавания.

Реализовать такой комплекс целесообразнее в составе НПУ, где для этого есть значительно большие возможности, а МБРЛС в рамках решения этой задачи должна выполнять требуемое информационное обеспечение, т.е. формирование детальных РЛИ.

Задача сопровождения цели начинает решаться после того, как оператор по уже полученному РЛИ осуществляет «сколку» объекта. В режимах картографирования или СНДЦ оператор выполняет захват выбранной в РЛИ цели, тем самым определяя положение строба по азимуту и дальности. Захват и сопровождение цели осуществляются по её энергетическому центру, положение которого оценивается либо по всем элементам строба (при работе по неподвижной цели), либо только по выделенным как движущиеся (при работе по мобильной цели). В режиме СНДЦ, где РЛИ не формируется, а на НПУ передаются только формуляры целей (координаты, скорость), метки обнаруженных движущихся целей накладываются на ранее полученное РЛИ или цифровую карту местности. С этого момента включается контур сопровождения и выполняется постоянное точное измерение координат и параметров движения объекта, которые в виде целеуказания выдаются разведывательным техническим или ударным средствам. Для обеспечения непрерывного функционирования контура сопровождения с НПУ на МБРЛС передаются текущие параметры зоны наблюдения, а МБРЛС непрерывно следит за целью в выбранном оператором режиме и возвращает на НПУ текущие значения координат и параметров движения объекта. Если это режим картографирования или СНДЦ с формированием РЛИ, то включается стабилизация центра зоны обзора. Все последующие РЛИ начинают формироваться с общим центром, при этом изображение по мере движения БЛА будет поворачиваться, как показано на рис. 12. В режиме СНДЦ без формирования РЛИ центр зоны обзора может быть «привязан» либо к заданной точке на поверхности (стабилизация), либо к энергетическому центру отметки от выбранной оператором цели (слежение).

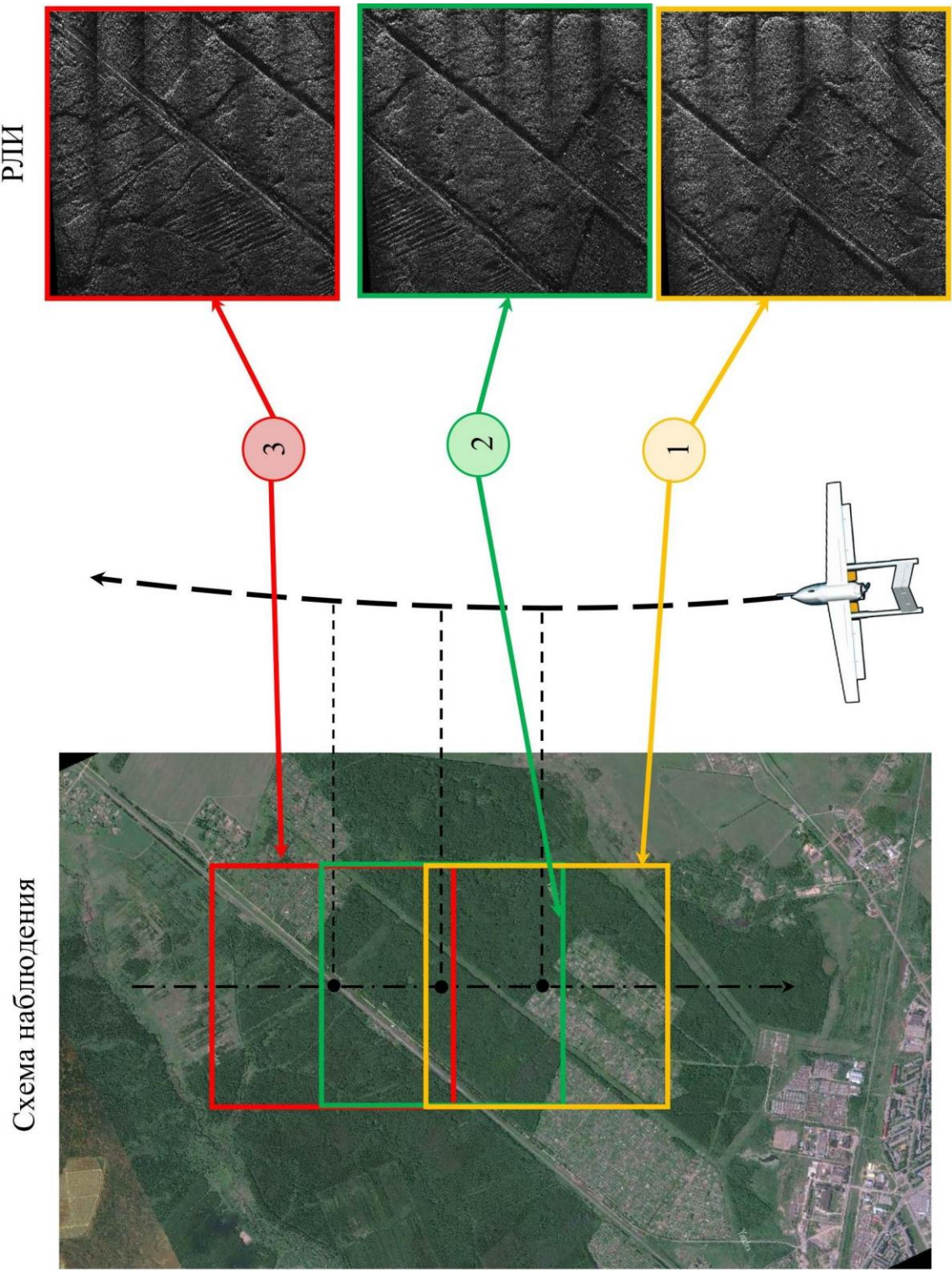
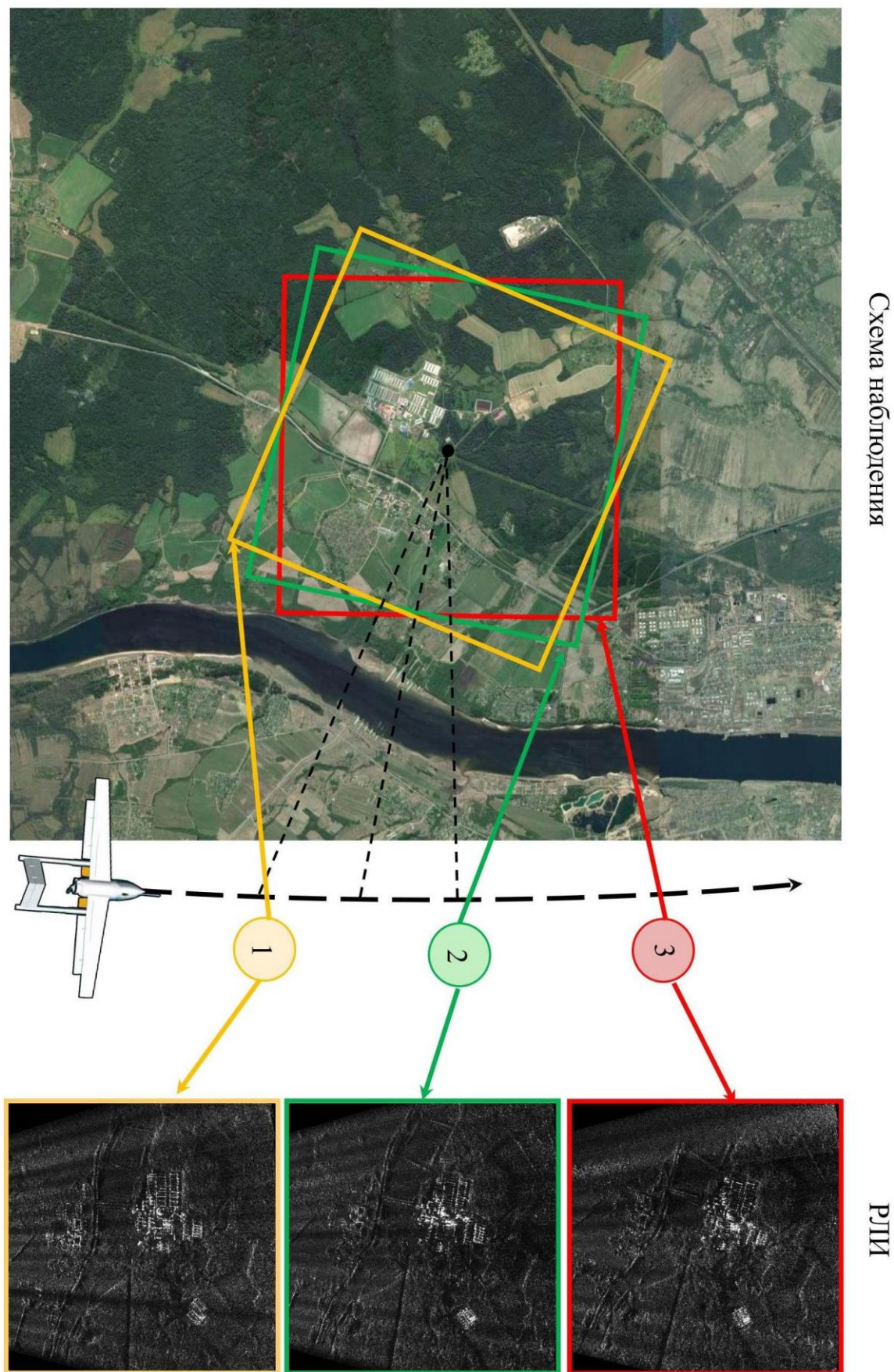


Рис. 11. Формирование РЛИ при мониторинге местности
(рисунок, выполненный в цвете, можно посмотреть на сайте журнала)



5. Заключение

Очевидно, что для оперативного решения всех перечисленных выше боевых задач режимы работы многофункциональной бортовой радиолокационной системы как элемента разведывательно-ударного комплекса с БЛА должны не только обеспечивать различные варианты функционирования (комбинации разрешений, углов наблюдения, видов обзора поверхности и т.д.), но и быть логически взаимосвязанными между собой, что реализовано в функциональном программном обеспечении МБРЛС.

Приложение

Список аббревиатур

БЛА МД и СД – беспилотный летательный аппарат малой и средней дальности.

ДЛГ – дальность до линии горизонта.

КБЛА – комплексы с беспилотными летательными аппаратами.

КРТ-СА – картографирование с синтезированием апертуры антенны.

КРТ-ПО – картографирование с полосовым обзором.

МБРЛС – многофункциональная бортовая радиолокационная система.

МРЛЦН – многофункциональная радиолокационная целевая нагрузка.

НПУ – наземный пункт управления.

ПК – парциальный кадр.

ПМН – подсистемы микронавигации.

РЛИ – радиолокационное изображение.

СНДЦ – селекция наземных (надводных) движущихся целей.

ТТХ – тактико-технические характеристики.

ФПО – функциональное программное обеспечение.

Литература

1. Ильин Е. М. и др. Малогабаритный многофункциональный бортовой РЛК для беспилотных летательных аппаратов малой дальности // Вестник СибГУТИ. 2017. № 4 (40). С. 104–109.
2. Гуськов Ю. Н., Самарин О. Ф., Савостьянов В. Ю. Многофункциональные малогабаритные бортовые РЛС. Критические технологии // Радиоэлектронные технологии. 2018. № 1. С. 40–46.
3. Брайткрайц С. Г. и др. Унифицированный интегрированный с подсистемой микронавигации малогабаритный многофункциональный радиолокатор для беспилотных летательных аппаратов средней и малой дальности // Известия ТулГУ. Технические науки. 2018. Вып. 11. С. 313–326.
4. Ильин Е. М., Полубехин А. И., Савостьянов В. Ю., Самарин О. Ф., Тростин Е. А. Многофункциональная малогабаритная РЛС Ku-диапазона для БЛА оперативно-тактического звена // Радиоэлектронные технологии. 2019. № 1. С. 36–42.

Статья поступила в редакцию 24.04.2019.

Ильин Евгений Михайлович

д.ф.-м.н., профессор, ведущий аналитик инновационного технологического центра комплекса научной политики МГТУ им. Н. Э. Баумана (105005, Москва, 2-я Бауманская ул., д. 5, стр. 1), тел. 8-910-433-27-89, e-mail: evgil45@mail.ru.

Репников Дмитрий Александрович

заместитель директора инновационного технологического центра МГТУ им. Н.Э. Баумана (105005, Москва, 2-я Бауманская ул., д. 5, стр. 1), тел. (499) 263-68-46, e-mail: bauman@bmstu.ru.

Савостьянов Владимир Юрьевич

к.т.н., доцент, начальник лаборатории ОАО «Корпорация «Фазотрон-НИИР» (123557, Москва, Электрический переулок, 1), тел. (495) 955-11-00, e-mail: info@phasotron.com.

Самарин Олег Федорович

к.т.н., с.н.с., начальник НИО ОАО «Корпорация «Фазотрон-НИИР» (123557, Москва, Электрический переулок, 1), тел. 8-916-447-94-06, e-mail: nio6.fazotron@yandex.ru.

Полубехин Александр Иванович

к.т.н., руководитель инновационного технологического центра комплекса научной политики МГТУ им. Н. Э. Баумана, тел. 8-499-263-68-46, e-mail: polub1980@mail.ru.

Черевко Александр Григорьевич

к.ф.-м.н., доцент, заведующий кафедрой физики, заведующий лабораторией физических основ телекоммуникаций СибГУТИ (630102, Новосибирск, ул. Кирова, 86), тел. 8-913-980-60-71, e-mail: perspl4@mail.ru.

The operation modes of airborne multifunctional radar of short- medium range UAVs

E. Il'in, D. Repnikov, V. Savost'yanov, O. Samarin, A. Polubekhin, A. Cherevko

The main operation modes of a multifunctional airborne radar system of UAVs short and medium range (MARS UAV) are considered. The developed tactics of application of MARS operation modes for detecting, identifying and tracking various objects as well as monitoring are presented. These tactics allow UAV operator to use effectively the technical capabilities of the MARS.

Keywords: Multi-functional airborne radar, radar target load, UAV, modes of operation, tactics, mapping; detecting, identifying and tracking objects.

Эволюционно-технологический подход в управлении требованиями к робототехническим комплексам военного назначения

С. И. Безденежных, С. Г. Брайткрайц, Д. А. Репников, В. И. Флегонтов,
А. Г. Черевко

Рассмотрен существующий механизм формирования и корректировки требований к робототехническим комплексам военного назначения в Минобороны России, предлагаются меры по его совершенствованию, в частности: переход от системы сбора и формирования требований к системе управления требованиями, основанной на эволюционно-технологическом подходе к развитию вооружения.

Ключевые слова: управление требованиями, техническое задание, эволюция технологий, робототехнические комплексы.

1. Введение

Обоснование требований к новому высокотехнологичному вооружению, военной и специальной технике (ВВСТ) является одним из центральных вопросов военной науки. Этой проблематике посвящены многочисленные исследования, которые затрагивают все виды ВВСТ, в том числе и робототехнические комплексы (РТК).

Наряду с этим постепенное усложнение систем вооружения привело к значительному росту количества предъявляемых к ним требований. Например, тактико-техническое задание (ТТЗ) на создание первой в СССР ядерной бомбы (утвержденное в 1946 г.) умещалось на одном листе, тогда как современные документы такого рода содержат десятки страниц.

Переход к конкурсной основе заключения контрактов на НИОКР с фиксированной ценой создал ряд ранее не существовавших проблем, таких как отсутствие оценки реализуемости предъявляемых требований, сложность изменения сроков и объема работ, которые могли бы компенсировать недостатки ТТЗ.

Быстрое развитие технологий приводит к увеличению темпов внедрения инноваций в разрабатываемые и модернизируемые образцы высокотехнологичной продукции. Однако сложность и затянутость процедур разработки и согласования ТТЗ, а также их проверки в ходе многочисленных испытаний приводят к тому, что скорость появления новых технологий превышает длительность цикла разработки комплексов РТК военного назначения. В итоге технологический уровень поступающих в войска РТК отстает от мирового.

Таким образом, проблемы создания новых РТК военного назначения связаны не столько с наличием научно-технического задела, сколько с организационными проблемами их разработки, особенно с управлением требованиями к новым изделиям.

Под *управлением требованиями* (requirements management) [1] понимают деятельность в области формирования требований к изделию, их структурирования и документирования,

взаимного согласования, проверки выполнения заданных требований, согласованного и контролируемого их изменения при необходимости.

Впервые потребность в системах управления требованиями возникла в начале 90-х годов в области разработки программного обеспечения (ПО) [2], когда число предъявляемых требований к ПО росло в геометрической прогрессии. Однако спустя десять лет эти технологии начали применяться для согласования требований к сложным изделиям машиностроения и даже стали частью дисциплины «Системная инженерия» [3].

Несмотря на то, что термин «управление требованиями» в военных кругах используется узким кругом специалистов, он уже нашел отражение в ГОСТ Р 56136-2014 «Управление жизненным циклом продукции военного назначения. Термины и определения». Как не сложно заметить, новое понятие объединяет процедуры, которые и ранее существовали в Минобороны России, однако не были систематизированы и автоматизированы.

2. Процесс управления требованиями в МО РФ

Существующий процесс управления требованиями в МО РФ начинается с определения облика Вооруженных Сил в концептуальных документах (рис. 1). Далее требования развиваются и отражаются в тематических карточках и справках обоснования на НИР (ОКР). При успешном включении соответствующей работы в Государственную программу вооружения начинается разработка ТТЗ, в котором аккумулируются требования документов системы общих технических требований, государственных стандартов, научно-технический задел, опыт эксплуатации и боевого применения аналогичных комплексов, опыт обучения на аналогичные комплексы, перспективы развития средств борьбы и др.

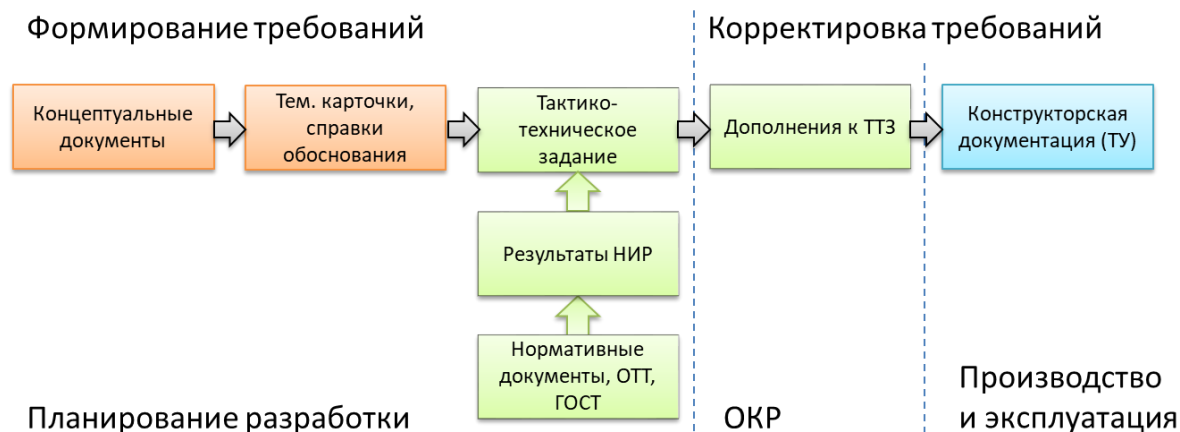


Рис. 1. Документы, отражающие требования к изделию в ходе ЖЦ

Существующая система управления требованиями допускает корректировку требований в ТТЗ, однако этот процесс сопряжен с рядом трудностей, таких как изменение существенных условий контракта и определение стоимости работ. Контроль выполнения требований на этапах выполнения ОКР производится комиссией заказчика экспертным методом. Главным этапом контроля требований по ГОСТ РВ 15.210-2001 является этап государственных испытаний, в ходе которого выполнение требований проверяется по согласованной и утвержденной программе и методикам. После проведения государственных испытаний контроль требований ТТЗ отсутствует.

В настоящий момент цель управления требованиями при создании РТК можно сформулировать как сбор и всестороннее обоснование тактико-технических требований для согласованного формирования ТТЗ, которое позволит с максимальной вероятностью создать изделие, реализующее заданные требования и имеющийся научно-технический задел.

Существующей системе присущ ряд недостатков: жесткое этапирование работ по стадиям жизненного цикла (ЖЦ), ориентированность на создание конкретного образца ВВСТ,

малый срок, выделенный на формирование требований, отсутствие учета НТЗ, сложность внесения изменений после заключения контракта, ориентация на выполнение требований ТТЗ, а не на создание изделия, отсутствие учета требований, выявленных после проведения государственных испытаний, разная глубина проработки процедур на разных этапах (некоторые частично не формализованы).

3. Модели развития вооружения

Во многом все упомянутые проблемы обусловлены используемым сегодня в Минобороны России комплексом стандартов СРПП ВТ. Этот комплекс отражает характерную для периода 1970–1990 гг. модель развития вооружения – так называемую *каскадную* (водопадную) модель [4], которая предусматривает последовательное выполнение всех этапов проекта в строго фиксированном порядке. Каждый этап завершается после полного выполнения и документального оформления всех предусмотренных работ.

В 1986 году Барри Бозом была предложена альтернатива каскадной модели – спиральная (*инкрементная*) модель процесса разработки [5]. По мнению многих она стала существенным прорывом в понимании природы разработки программного обеспечения. При таком подходе проектные задачи группируются в виде фаз (витков спирали), направленных на разработку последовательности поставляемых изделий, каждое из которых со временем наращивает свои функциональные возможности. Первая конструкция реализует часть требований, в последующую конструкцию добавляют дополнительные требования и так далее до тех пор, пока не будет закончено создание системы.

Вскоре инкрементная модель трансформировалась в *эволюционную модель*. Она также предполагает разработку последовательности конструкций, но, в отличие от инкрементной модели, подразумевается, что требования не могут быть полностью осознаны и сформулированы предварительно, поэтому они устанавливаются частично и уточняются в каждой последующей конструкции. Процессы сопровождения и эксплуатации могут быть реализованы параллельно с процессом разработки. Процессы заказа и поставки, а также вспомогательные и организационные процессы обычно выполняют параллельно с процессом разработки.

В середине 90-х годов эволюционная модель разработки начала набирать популярность не только в среде информационных технологий, но и в других областях разработки высокотехнологичной продукции. В 2003 году наряду с каскадной моделью эволюционная модель под именем «Эволюционное приобретение» (Evolutionary Acquisition) вошла в Инструкцию МО США №5000.02 «Функционирование системы оборонного заказа» – это главный документ МО США, определяющий порядок разработки ВВСТ [6]. При очередном пересмотре Инструкции №5000.02 в 2008 году эволюционный заказ стал предпочтительным способом создания ВВСТ в Минобороны США.

Таким образом, например, ВВС США поступили при создании комплекса с БпЛА GlobalHawk [7]. БпЛА первого этапа, сборка которых началась в 2004 г., были выпущены с использованием технологий опытного образца 2002 г. Сборка образцов второй итерации в 2005 г. была запланирована с использованием новой полезной нагрузки и внутренних компонентов. К моменту начала сборки БпЛА третьей итерации в 2006 г. было запланировано окончание работ по доработке двигателя, крыла и линии связи.

Наряду с этим еще в начале XX века ученые-экономисты обратили внимание на то, что процесс развития техники и технологий имеет определенное сходство с теорией биологической эволюции Дарвина. На эволюционное развитие идей при создании технических систем указывали авторы теорий решения изобретательских задач, поискового конструирования, адаптации сложных систем. Подобным образом ряд исследователей подтверждали эволюционный характер развития вооружения.

В 2009 г. известный американский экономист Б. Артур впервые подробно рассмотрел эволюцию сложных технических систем с точки зрения развития составляющих их техноло-

гий [8]. Незадолго до этой фундаментальной работы главенствующую роль эволюции технологий в развитии техники и вооружения отмечали отечественные военные ученые [9, 10].

В результате был предложен новый эволюционно-технологический подход (ЭТП) к развитию вооружения – методология создания и приобретения ВВСТ, основанная на постепенном многошаговом процессе повышения уровня знаний в определенных научно-технологических областях [11]. Идея, лежащая в основе ЭТП, состоит в том, что образец следует разрабатывать по принципу приращений, так, чтобы разработчик мог использовать данные и знания, полученные при разработке более ранних версий изделия или схожих систем.

ЭТП рассматривает технологию как знание. Жизненный цикл технологии представляется эволюцией знания. От идеи до реализации изделие происходит итерационный процесс дополнения и проверки достоверности знания, закреплённой в модели изделия. То есть процесс разработки является процессом дополнения модели изделия достоверными знаниями.

Роль эволюции технологий в ЭТП заключается в том, что переносчиками наследственной информации изделий являются составляющие их технологии. Так же как гены в биологических системах, технологии технических систем, скрещиваясь, обмениваются решениями. Выживают лишь наиболее лучшие виды техники, наиболее приспособленные к среде обитания – потребностям вооруженной борьбы. При этом подразумевается, что требования не могут быть полностью осознаны и сформулированы предварительно, поэтому они устанавливаются частично и уточняются в каждой последующей конструкции.

В работе [12] предложена оригинальная модель эволюции технологий в функциональном, технологическом и физическом доменах. Ее суть заключается в том, что описание любой технологии можно представить в виде структурных моделей: функциональной, технологической и физической (рис. 2). Эволюция происходит последовательным изменением этих моделей, главенствующей из которых является технологическая.

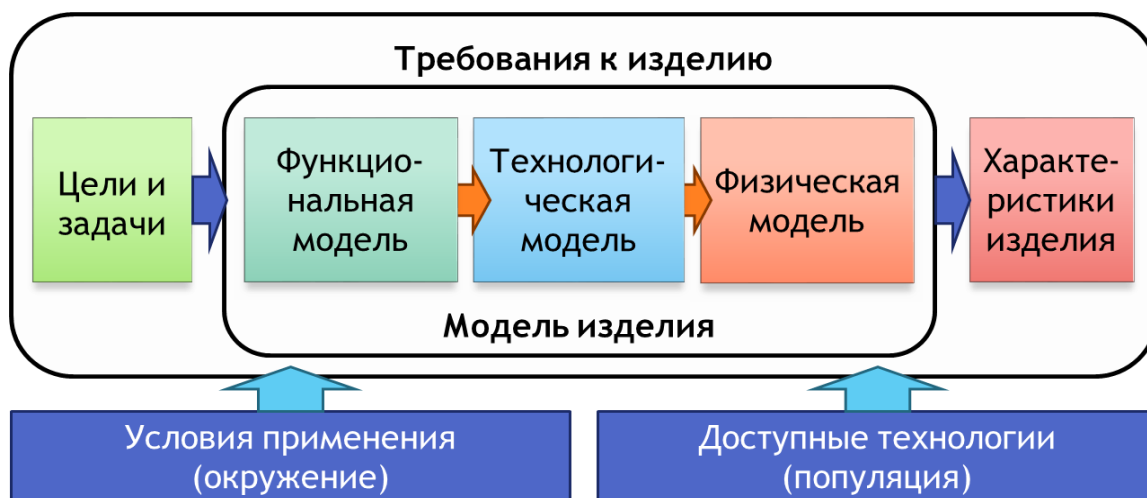


Рис. 2. Структура требований к изделию в соответствии с ЭТП

При таком подходе требования рассматриваются как часть модели изделия, а также как сведения об условиях применения и необходимых характеристиках. Особый акцент делается на то, что требования отражают не достоверное знание, а представление потребителей о том, как должно выглядеть изделие для того, чтобы эффективно решать свои задачи.

4. Эволюционно-технологический подход к управлению требованиями

В части управления требованиями обычный эволюционный подход провозглашает переход от разового, жесткого, неизменяемого набора требований к постоянному процессу уточнения технического задания и порождению параллельного потока работ по созданию

изделия. При этом акцент делается не на выполнение всех требований ТТЗ, а на создание и представление заказчику изделия, имеющего полезные характеристики (свойства), способного заменить неудовлетворяющие современным требованиям стоящие на вооружении образцы ВВСТ.

Эволюционно-технологический подход порождает еще несколько отличий от традиционного подхода к управлению требованиями.

Во-первых, изначально предполагается, что представления потребителей могут быть ошибочны. Вследствие чего жесткая фиксация перечня требований исключается. Ошибка может быть выявлена как на этапе разработки и привести к невозможности создания изделия, так и на этапе эксплуатации и снизить ожидаемую эффективность.

Во-вторых, требования делятся на пять структурированных множеств: назначения, функциональные, технологические, физические и характеристики. Требования верхнего уровня имеют приоритет перед остальными.

В-третьих, так как предполагается наличие постоянного процесса изменения требований, производится мониторинг популяции технологий и окружения с целью выявления изменений в моделях технологии.

В-четвертых, эволюционная природа модели предполагает явное использование прототипов. Это относится как к этапам разработки требований для вариантов изделия, что в принципе характерно для систем управления требованиями, так и к этапу первоначального формирования требований, когда структура изделия наследуется от одного или нескольких ранее созданных изделий (прототипов, опытных образцов).

Ключевыми особенностями предлагаемого подхода к управлению требованиями являются следующие:

1. Требования рассматриваются как представление заказчика о назначении, функциях, составе и параметрах изделия, обладающего достаточной эффективностью. Учитывается тот факт, что представления могут быть ошибочны как с точки зрения эффективности, так и с точки зрения возможности реализации, и требования потребуют изменений.

2. Из состава требований выделяются назначение и задачи изделия. Главным критерием при принятии изделия на вооружения (снабжение) становится выполнение изделием своих задач, а не соответствие всем условиям технического задания.

3. Управление требованиями происходит не в рамках создания отдельного образца вооружения (программы), а в ретроспективе развития типажа вооружения.

4. С использованием структурных функционально-технологических схем происходит систематический мониторинг возможностей аналогичных комплексов (популяции), состояния базовых и критических технологий, а также условий функционирования для каждого типа вооружения.

Структурные функционально-технологические схемы являются основой для выбора аналогов и формирования новых тактико-технических заданий на разработку и модернизацию новых изделий

5. Заключение

Изложенный подход к управлению требованиями объединяет достоинства современных технологий управления требованиями и эволюционно-технологического подхода. Результатом его применения должно стать сокращение рисков заказчика от ошибок в ТТЗ, сокращение сроков получения полезного образца и ускорение внедрения в систему вооружения нововведений. Однако полноценное внедрение ЭТП к управлению требованиями может быть реализовано только после проработки соответствующего научно-методического аппарата и создания специализированных программных средств.

Литература

1. ГОСТ Р 56136-2014 «Управление жизненным циклом продукции военного назначения. Термины и определения».
2. Davis A. M. Just Enough Requirements Management: Where Software Development Meets Marketing. Dorset House, 2005. 240 с.
3. Косяков А., Свит У., Сеймур С, Бимер С. Системная инженерия. Принципы и практика / пер. с англ. В. Батоврин. М.: ДМК Пресс, 2014. 636 с.
4. Буренок В. М., Косенко А. А., Лавринов Г. А. Техническое оснащение Вооруженных Сил Российской Федерации: организационные, экономические и методологические аспекты. М.: Граница, 2007. 728 с.
5. B. Boehm. Spiral Development: Experience, Principles, and Refinements // University of Southern California, 2000. URL: <http://csse.usc.edu/TECHRPTS/2000/usccse2000-507/usccse2000-507.pdf> (дата обращения: 20.02.2019).
6. Безденежных С. И., Артеменко В. Б. Обзор системы оборонного заказа МО США // Вооружение и экономика. 2014. № 1 (26).
7. Shah N. B. Modularity as an Enabler for Evolutionary Acquisition. // Massachusetts institute of technology. 2004. URL: http://seari.mit.edu/documents/theses/SM_SHAH.pdf (дата обращения: 20.02.2019).
8. Arthur W. B. The Nature of Technology: What it Is and How it Evolves. The Free Press, and Penguin (UK) August 11, 2009. 246 p.
9. Буренок В. М., Ивлев А. А., Корчак В. Ю. Эволюционно-технологический подход к созданию перспективного вооружения // Военный парад. 2006. № 5–6;
10. Безденежных С. И. Эволюционно-технологический подход к развитию ВВТ // ЦВНИ МО РФ, 2006.
11. Буренок В. М., Ивлев А. А., Корчак В. Ю. Развитие технологий XXI века: проблемы планирование, реализация. Тверь: ООО «Купол», 2009. 624 с.
12. Безденежных С. И., Ивлев А. А. Формальная модель эволюции системных технологий в функциональном, технологическом и физическом доменах. // Электронный научный журнал «Исследовано в России». 2013. URL: <http://zhurnal.ape.relarn.ru/articles/2013/016.pdf> (дата обращения: 17.08.2016).

Статья поступила в редакцию 24.04.2019.

Безденежных Сергей Игоревич

заместитель начальника исследовательского отдела Государственного центра беспилотной авиации (140415, Московская обл., Коломна, пр-д Артиллеристов, 5), тел. (496) 618-55-58.

Брайткрайц Сергей Гарриевич

д.т.н., с.н.с., ведущий научный сотрудник 46 ЦНИИ МО РФ (129327, Москва, Чукотский пр-д. вл.10), тел. (495) 471-27-90.

Репников Дмитрий Александрович

заместитель директора инновационного технологического центра МГТУ им. Н. Э. Баумана (105005, Москва, 2-я Бауманская ул., д. 5, стр. 1), тел. (499) 263-68-46, e-mail: bauman@bmstu.ru.

Флегонтов Виталий Иванович

к.э.н, доцент к.э.н, доцент кафедры экономики и финансов Одинцовского филиала МГИМО МИД РФ (105005, Москва, 2-я Бауманская ул., д. 5), тел. (499) 263-68-46, e-mail: Fvi55@yandex.ru.

Черевко Александр Григорьевич

к.ф.-м.н., доцент, заведующий кафедрой физики, исп. директор Научно-образовательного центра «Материаловедение и микрофотоэлектроника» СибГУТИ (630102, Новосибирск, ул. Кирова, 86), тел. 8-913-980-60-71, e-mail: persp14@mail.ru.

Evolutionary and technological approach to managing requirements for military robotic complexes

S. Bezdenzhnykh, S. Braytkrayts, D. Repnikov, V. Flegontov, A. Cherevko

This paper considers the existing mechanism of formation and adjustment requirements for robotic complexes in the Ministry of Defense of Russia. We propose the ways to improve it: the transition from the "collection of requirements" system to the "management of requirements" system based on an evolutionary and technological approach.

Keywords: requirements, management, technical task, evolution of technology, robotic complexes.

Современные технологии в системе ММО

В. П. Петров, И. Ю. Якушев

В статье рассматриваются модели ММО с различными вариантами технологий для оптимизации характеристик системы в условиях неселективного канала Релея. Основное внимание уделено сравнительно новым антенным технологиям, которые рекомендуются для сетей пятого поколения – серии технологий пространственной модуляции и технологии антенной селекции. Приведены расчетные формулы производительности и емкости канала и графики на их основе для сравнения характеристик.

Ключевые слова: 5G, BER, ММО, канал Релея, емкость канала, антенные технологии.

1. Введение

Эволюция систем мобильной связи в последние 20 лет была направлена на увеличение производительности и уменьшение вероятности ошибки в городских каналах «передатчик – приемник». Основная сложность в достижении положительных результатов состояла в неопределенности характеристик канала. Функции плотности вероятности Релея, Райса и их обобщений Накагами не снимали неопределенности в моментах распределений, создавая проблему «связь над обобщенными fading-каналами (с замираниями)» [1]. Методология решения комплекса задач, казалось бы, лежит на поверхности. Это представление канала в виде четырехполюсника и известные методы решения с заданными требованиями к каналу. «Заданные требования» могут быть определены либо теоретической оценкой, либо измерениями в процессе организации связи. И то, и другое – с неточностями, которые проявляются в ошибках передачи.

Известно правило в теории передачи сигналов: чем больше сигнальной информации на входе/выходе 4-х полюсника, тем больше информации можно получить о его параметрах. Это правило реализуется в разнесении, в первую очередь в пространственном разнесении антенн на приемнике и на передатчике. Разнесение только на приемнике (система SIMO – *Simple-input-Multiple-output*) позволяет создать множество различных копий на элементах приемной антенной решетки и определить переданный сигнал комбинированием по многим, не совпадающим между собой, разнесенным искаженным сигналам [1, 5]. В системе MISO (*Multiple-input-Simple-output*) с разнесением антенн на передатчике имеются разнообразные возможности, связанные с формированием сигналов и их распределением по элементам антенной решетки. Применительно к системе ММО (*Multiple-input-Multiple-output*) эти возможности реализуются совместно с приемником в различных вариантах, создавая комплекс антенных технологий.

В статье, не претендуя на полноту анализа, рассматриваются перспективные технологии радиосвязи, в основу которых положена система ММО.

2. Радиоканалы и их роль в антенных технологиях

Информация о характеристиках радиоканала между передатчиком и приемником является важным фактором в эффективности системы ММО. Многолучевые каналы, характерные в городских условиях распространения радиоволн, неизбежно порождают задержку сигналов на элементах приемной антенной решетки, что, в свою очередь, потенциально порождает межсимвольную интерференцию (*ISI – intersymbol interference*).

Выбор способа обработки сигналов в приемнике, исключающего появление *ISI*, зависит от свойств рассеивающей среды распространения сигналов. При оценке влияния канала на искажение сигнала особое значение имеет сопоставление частотно-временных характеристик сигнала с характеристиками задержки на приеме. Действительно, лучи могут формироваться близко расположенными препятствиями, при которых разность хода, а значит, и задержка малы, и другой вариант – отражение от далеко расположенных препятствий с большими задержками на приеме. В канале с замираниями взаимосвязь между максимальной задержкой T_m и временем передачи символа T_s приводят к различным классификационным признакам канала в частотной области:

1) *частотно-независимый (плоский) канал*, когда время задержки мало ($T_m < T_s$), тогда полоса когерентности канала B_m больше полосы частот сигнала B_s ; и

2) *частотно-селективный канал*, когда $T_m > T_s$, а, следовательно, полоса частот сигнала B_s больше полосы когерентности канала B_m .

Полоса когерентности канала B_m – это статистическая мера диапазона частот, в рамках которого канал пропускает все спектральные компоненты сигнала с приблизительно равным коэффициентом ослабления и линейным изменением фазы. В данном случае (плоский канал) не порождает межсимвольную интерференцию *ISI*, так как расширение сигнала во времени не приводит к наложению соседних символов. В частотно-селективном канале многолучевой компонент символа выходит за пределы длительности передачи символа и порождает межсимвольную интерференцию *ISI*.

Сети и каналы с плотной городской застройкой, включая каналы внутри помещений, относятся к плоским каналам при передаче сигналов с узкополосной модуляцией, к которым относятся М-арные сигналы. Однако эти же каналы при передаче широкополосных сигналов OFDM, WCDMA относятся к частотно-селективным.

При анализе в приемнике плоских каналов ММО наибольшее применение получили следующие три метода:

- Комбинирование с автовыбором (*Selection Combining – SC*);
- Комбинирование с равным усилением (*Equal Gain Combining – EGC*);
- Комбинирование с максимальным отношением (*Maximal Ratio Combining – MRC*).

Первый из этих методов (*SC*) оценивает отношение сигнал/шум со всех антенных элементов и выбирает наибольшее значение, которое принимается за решение. Метод реализуется относительно просто, но он не является оптимальным, поскольку в нем не используются все полученные сигналы.

Метод *EGC* предполагает равными модули коэффициентов передачи лучей между передатчиком и приемником. Поэтому весовые коэффициенты элементарных сигналов равны единице. Метод также прост в реализации, но условие является необоснованным.

В методе *MRC* элементарные сигналы взвешиваются согласно значению отношения сигнал/шум (ОСШ) на каждом антенном элементе и когерентно суммируются. Информация с антенных элементов в этом случае используется наиболее полно и метод считается наиболее достоверным.

В системе ММО с разнесением антенн на передатчике имеются разнообразные возможности, связанные с формированием сигналов и их распределением по элементам антенной решетки. Применительно к системе ММО эти возможности реализуются совместно с приемником в различных вариантах, создавая комплекс антенных технологий. При этом важную роль

играет наличие информации на передатчике о характеристиках канала. Возможны два варианта: информация на передатчике отсутствует – *открытый канал*, и информация на передатчике доступна – *закрытый канал*. Канал называется закрытым, поскольку для получения информации должен присутствовать канал обратной связи. Этот канал не требует высокой скорости, так как требуемое количество информации о канале значительно уступает сигнальному объему. В обоих случаях характеристики канала известны приемнику путем анализа тестового сигнала передатчика.

Заметим также, что следует различать системы MIMO: система «точка-точка», или однопользовательская система (*Single User MIMO – SU-MIMO*), и система «точка-многоточка», или многопользовательская система (*Multi User MIMO – MU-MIMO*). Названия систем отражают их назначение. Первая – канальная система, вторая – сетевая. В последнее время в связи с предстоящим внедрением 5G появился термин «*Massive MIMO*», который относится к сетевому виду, но с очень большим числом антенных элементов в решетке, создающим особенности анализа.

Ниже будут рассмотрены канальные системы MIMO применительно к частотно-независимым каналам. Их математические модели и анализ являются базовыми для применения в других системах. Задачей анализа является поиск оптимальных условий реализации MIMO, обеспечивающих наибольшую емкость системы и требуемый уровень помехоустойчивости.

3. Обзор антенных технологий системы MIMO

Действующие антенные технологии канальных систем MIMO разнообразием не отличаются и во всех случаях канальная модель в условиях частотно-независимого многолучевого канала в матричной форме описывается уравнением:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

где $\mathbf{y}$ есть матрица-столбец сигналов приёмника, описывающая ансамбль сигналов на элементах приёмной решётки;

$\mathbf{x}$ – матрица-столбец сигналов передатчика, описывающая ансамбль сигналов, излучённых каждым из элементов передающей решётки;

$\mathbf{n}$ – матрица-столбец аддитивных шумов приёмника и помех, создаваемых рассеивающей средой в пространстве между передатчиком и приёмником;

$\mathbf{H}$ – прямоугольная комплексная матрица канала размерности $M \times N$. Компонент h_{ij} этой матрицы есть комплексный коэффициент передачи канала между j -ой передающей антенной и i -ой приёмной антенной. При этом M – количество антенн в приёмной решетке, N – количество антенн в передающей решетке.

В условиях замираний элементы матрицы канала являются случайными величинами, равными относительно главной диагонали ($h_{ij} = h_{ji}$). Как следует из (1), оценка принятого сигнала $\mathbf{y}$ более достоверна, если известна канальная матрица $\mathbf{H}$.

На рис. 1а представлена схема лучей, связывающая передающую и приёмную антенные решетки.

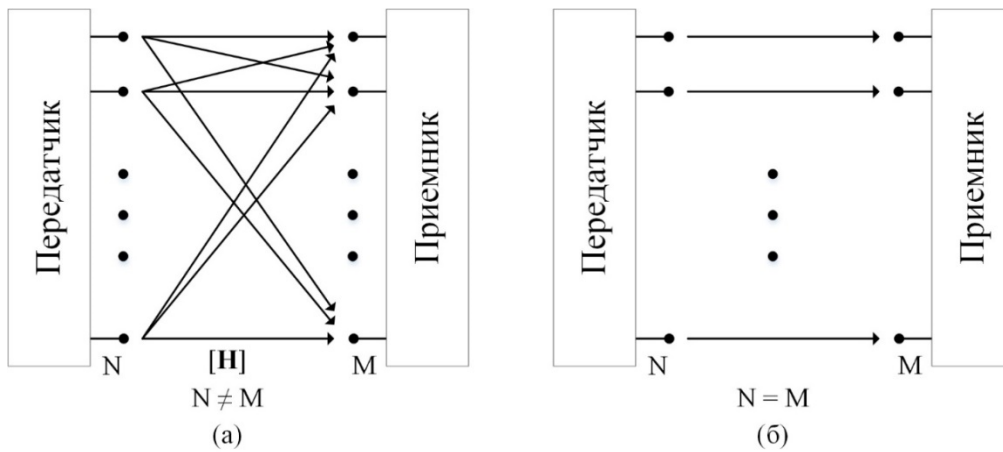


Рис. 1. Система ММО

Преобразование сложной системы потоков в виртуальную систему параллельных потоков от передатчика к приемнику (рис. 1б) осуществляется известным из теории матриц [2, 8] *сингулярным* разложением матрицы $\mathbf{H}$ в виде:

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H. \quad (2)$$

В этом уравнении $\mathbf{U}$ и $\mathbf{V}$ – унитарные матрицы с размерами $M \times M$ и $N \times N$ соответственно, $\mathbf{U}^H\mathbf{U} = \mathbf{I}_M$ и $\mathbf{V}^H\mathbf{V} = \mathbf{I}_N$, и $\mathbf{\Sigma}$ – есть диагональная матрица сингулярных значений d_j матрицы $\mathbf{H}$ размером $M \times N$. Известно, что сингулярные значения d_j и собственные значения λ_j связаны соотношением:

$$d_j = \sqrt{\lambda_j}.$$

Количество антенных элементов на передатчике и приемнике, как правило, различно ($M \neq N$). Пример – базовая и мобильная станции. Излучение N антенн распределяется на M приемных антенн. Чтобы представить реальный канал в виде виртуальных параллельных каналов SISO, необходимо допустить равенство антенн передатчика и приемника. Этим равенством является минимальное число M или N , которое в математических моделях является рангом канальной матрицы. В результате преобразований показано [2], что каждый элемент вектора $\mathbf{y}$ может быть представлен следующим образом:

$$y_i = \sqrt{\lambda} \cdot x_i + [\mathbf{U}^H \mathbf{n}]_i. \quad (3)$$

Это уравнение характерно тем, что имеет только соответствующие i -ые лучи от передатчика к приемнику. Отсутствуют канальные лучи h_{ij} . То есть ММО каналы преобразованы в виртуальные параллельные SISO каналы, как отражено на рис. 1б.

Когда информация о канале отсутствует на передатчике, мощность передатчика обычно равномерно распределяется по излучающим элементам антенной решетки. Однако, когда информация известна передатчику, мощность передатчика может быть оптимально распределена по излучающим некоррелированным элементам с целью максимального увеличения спектральной эффективности. Этот способ называется *алгоритмом «заполнения водой»* (WFA – *water-filling algorithm*) [1, 2]. Алгоритм позволяет сконцентрировать излучаемую мощность только в «хороших» пространственных каналах и не тратить энергетические ресурсы передатчика на попытки передать информацию через «плохие» пространственные каналы. «Хорошие» каналы – это каналы, имеющие высокое ОСШ, «плохие» – наоборот. Возможность оценки качества каналов лежит в основе новых технологий.

В результате применения этого алгоритма оптимальная спектральная эффективность определяется формулой [2]:

$$\eta = \sum_{i=1}^N \log_2 \left(1 + \frac{P_i}{\sigma^2} \lambda_i \right) = \sum_{i=1}^N \log_2 \left(\left(\bar{\gamma} + \sum_{l=1}^N \frac{1}{\lambda_l} \right) \frac{\lambda_i}{N} \right). \quad (4)$$

На рис. 2 и 3 видно, что емкость системы MIMO повышается при использовании алгоритма WFA. Средняя емкость системы MIMO – это емкость канала (иногда называют эргодической емкостью), гарантированная емкость задается гарантированной вероятностью ошибки P_{out} . В случае, когда алгоритм WFA не используется, мощность излучается одинаково через все пространственные каналы. Построение графиков производится в MATLAB на основе написанных авторами М-функций.

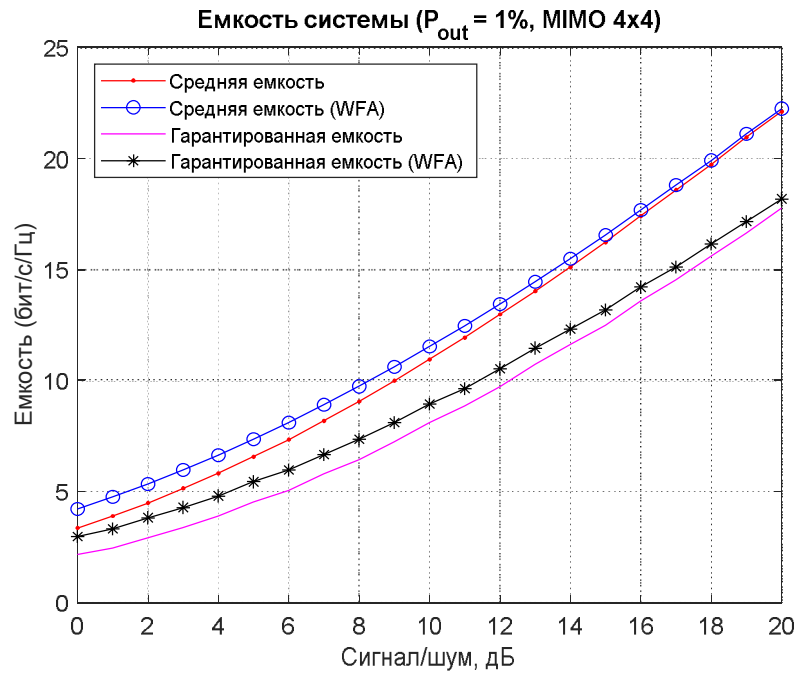


Рис. 2. Емкость системы MIMO с алгоритмом WFA

Рис. 3 отражает преимущество алгоритма WFA для гарантированной емкости системы MIMO 4x4 в зависимости от гарантированной вероятности ошибки при неизменном ОСШ.

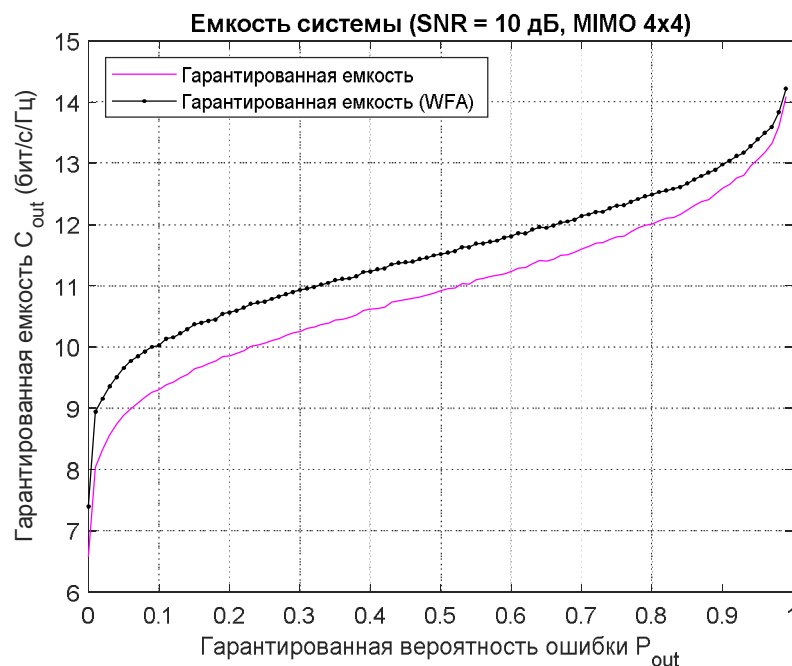


Рис. 3. Гарантированная емкость системы MIMO от P_{out}

Разложение на параллельные виртуальные потоки – это одна из стратегий увеличения емкости системы ММО. Антенные технологии в ММО-системах можно разделить на два вида (рис. 4): традиционные методы, включающие *пространственное мультиплексирование* (рис. 4а) и *пространственно-временное кодирование* (рис. 4б), и новые методы, включающие *пространственную модуляцию* (рис. 4в) и *селекцию передающих антенн* (рис. 4г). Рисунок дает упрощенное представление технологий с использованием двух антенн и двух сигнальных символов.

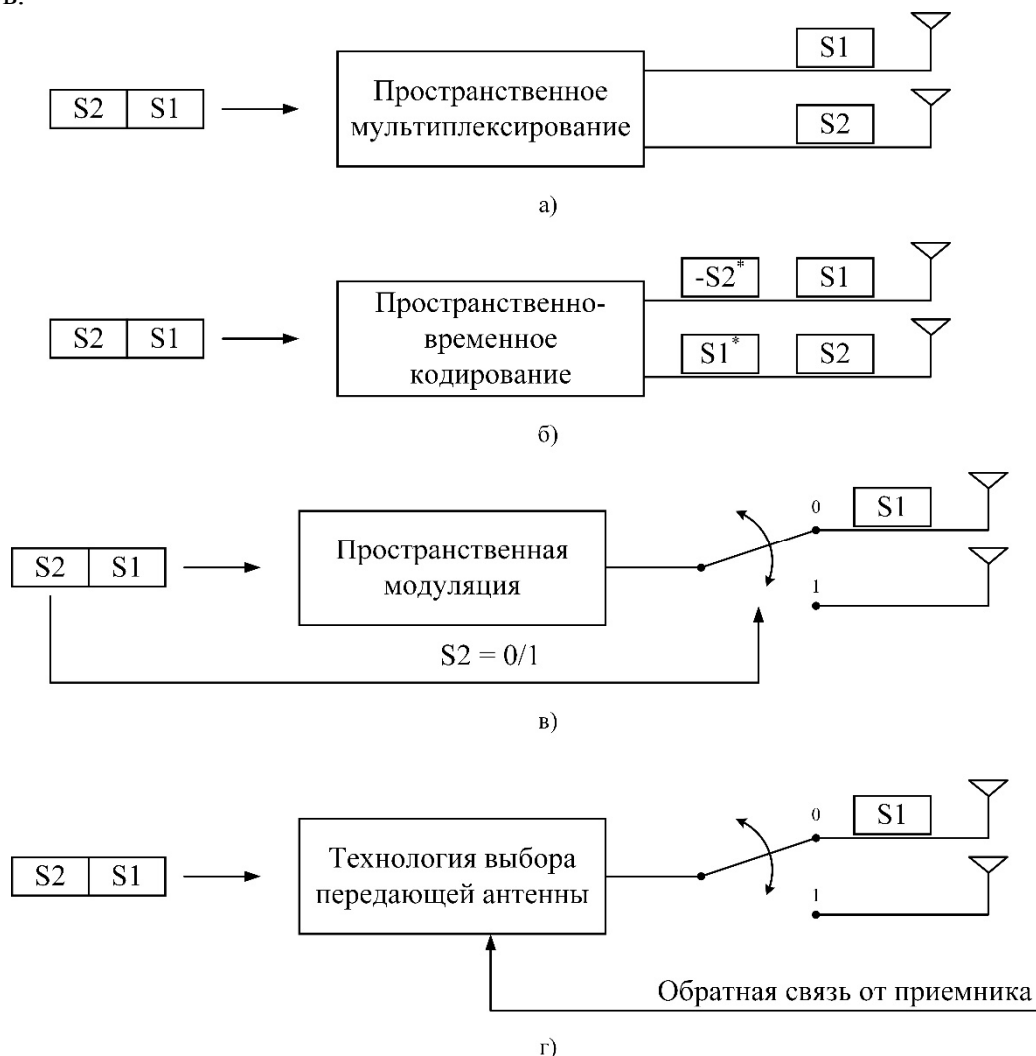


Рис. 4. Методы антенных технологий

При *пространственном мультиплексировании* (*Spatial Multiplexing – SMX*) два символа одновременно излучаются в одном временном слоте. В общем случае поток входящих данных преобразуется из последовательного в параллельный для передачи. Параллельные потоки информации, полученные после этого преобразования, передаются одновременно с нескольких антенн, доступных на передатчике, что обеспечивает высокую спектральную эффективность системы, поскольку скорость передачи данных становится пропорциональной числу передающих антенн [2, 5]. На этом принципе построена схема *BLAST* (*Bell Laboratory Space Time*).

В случае *пространственно-временного кодирования* (*Space Time Coding – STC*), например, схема Аламоути, два ортогональных символа одновременно излучаются в два временных слота. Эта схема хорошо известна и нашла применение в различных системах мобильной связи по причине простой реализации. Недостаток схемы состоит в том, что она реализуется только при двух передающих антеннах и, хотя может использовать несколько приемных антенн, тем не менее, спектральная эффективность схемы недостаточно велика.

Наибольший интерес представляют новые, наименее изученные, направления при исследовании производительности и помехоустойчивости сетей ММО в городских условиях. Это

пространственная модуляция сигналов (Spatial Modulation – SM) и технология выбора передающей антенны в последовательности излучателей сигналов (Transmit Antenna Selection – TAS).

Новые технологии требуют более детальных пояснений. Они родились и были частично исследованы в рамках четвертого поколения [1–5]. Особый интерес проявился к ним в связи с разработкой технологий пятого поколения. Интерес обусловлен не только потенциальным увеличением емкости канала, но и существенным увеличением энергетической эффективности технологий [6].

4. Характеристики технологий пространственной модуляции

Применительно к двум антеннам (рис. 4в) *пространственная модуляция* работает просто: бит 1 – излучает активная антенна, бит 0 – другая антенна не излучает. При передаче последовательного потока из единиц и нулей антенны переключаются, как клавиши на пианино. Предположим теперь, что передающих антенн множество и каждая антенна передает определенный поток битов – набор звуков, присущих данной клавише (антенне). Прием при таком режиме передачи заключается только в извлечении из банка данных передатчика, известного приемнику, последовательных индексов активных антенн. Обратим внимание, что в этом режиме сигнальные биты не передаются физически. При двух антеннах выигрыш мал, не более 3 дБ, но чем больше число передающих антенн, тем разнообразнее варианты и тем больше выигрыш в производительности и в помехоустойчивости при малых затратах энергии, т.е. при высокой энергетической эффективности.

Расширенная блок-схема формирования сигнала в передатчике пространственной модуляции представлена на рис. 5, где блок S/P – преобразование последовательности бит данных в параллельный поток, ЦАП – цифро-аналоговый преобразователь, RF Switch – блок выбора активной антенны.

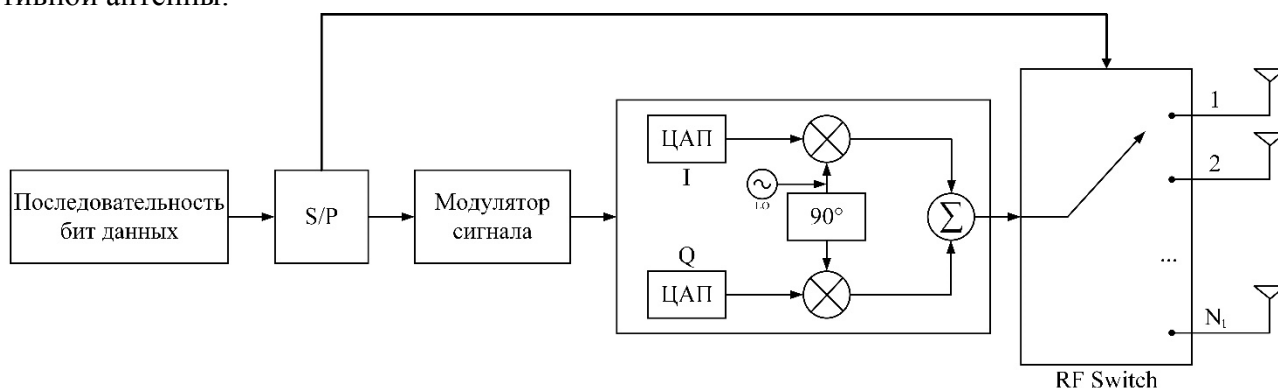


Рис. 5. Блок-схема пространственной модуляции

Предположим, что из N_t передающих антенн в решетке n_u различных антенн являются активными, причем каждая активная антенна имеет свой кодовый набор нулей и единиц. Число комбинаций из N_t по n_u определяется соотношением:

$$C_{N_t}^{n_u} = \frac{N_t!}{n_u!(N_t - n_u)!}. \quad (5)$$

Так, например, $C_4^2 = 6$; $C_6^3 = 20$; $C_8^4 = 70$. Число комбинаций стремительно растет с ростом числа передающих антенн, часть из которых является активными. Этот фактор, как следует из теории пространственной модуляции [7, 8], приводит к выигрышу основных характеристик системы ММО, в том числе *энергетической эффективности* системы. Этот важный параметр имеет тенденцию уменьшаться с ростом *спектральной эффективности* при традиционных

технологиях ММО, но не в данной схеме. Пространственная модуляция особенно эффективна в миллиметровом диапазоне, как планируют в 5G, при антеннах из многих десятков излучателей, поддерживая 4 – 5 активных.

Почему эта технология называется *пространственной модуляцией SM*? Это название она получила, дополняя плоскую звездную диаграмму комплексного сигнала пространственной координатой активных излучателей (индексы антенн), что создает трехмерную пространственную диаграмму [3, 4, 5]. В SM все биты информации не передаются физически. Часть битовой последовательности информации отображается для передачи индексов антенны (*пространственное созвездие*), и антенна, соответствующая отображенному индексу, используется для передачи информационных битов, отображенных на созвездие сигнала. Модель SM ММО рассматривает систему $N_t \times N_r$ ММО, имеющую N_t антенн на стороне передатчика и N_r антенн на стороне приемника. На приемной стороне проводится оптимальная комбинаторная обработка сигналов. Для M-арной схемы модуляции каждый блок из $\log_2(N_t) + \log_2(M)$ битов информации сопоставляется в SM для созвездия сигнала и пространственных созвездий. Биты $\log_2(N_t)$ каждого блока отображаются на точки в пространственных созвездиях, которые выбирают антенну, а биты $\log_2(M)$ отображаются в созвездии сигнала и модулируются модулятором символов с использованием подходящей схемы цифровой модуляции. Затем точка созвездия цифрового модулированного сигнала передается от антенны, на которую сопоставляется точка пространственного созвездия.

Таким образом, используя SM, мы эффективно получаем спектральный коэффициент полезного действия $\log_2(N_t)$ бит/с/Гц без каких-либо дополнительных затрат полосы пропускания или мощности.

Метод не требует знания характеристик канала, то есть нет необходимости в обратной связи с приемником, что является положительным фактором при «on-line» связи.

В сущности, пространственная модуляция SM – это комплекс антенных технологий под общим названием SMTs (*Space Modulation Techniques – технологии пространственной модуляции*), объединенных общим правилом: технологии используют антенные индексы, при которых сигнальная информация не передается физически, а извлекается из банка данных. В крайнем случае, который рассмотрен выше, передача осуществляется только в режиме переключения индексов антенн (*SSK – Space Shift Keying*). При этом можно переключать последовательно либо по одной антенне (*SSK*), либо по две и более – обобщенный режим, или *GSSK*. Очевидно, что второй вариант более производителен. Оба варианта удобно использовать при передаче бинарных сигналов.

SSK – простейшая форма семейства SMTs. В *SSK*-системе данные передаются только через пространственный символ, а переданный сигнал есть немодулированный радиочастотный сигнал, отображающий пространственный индекс активной антенны. Антенны питаются обычно малой выходной мощностью и увеличивают мощность только при передаче сигнала с антенны.

На рис. 6 приведено сравнение большинства типов *SSK*-модуляции для системы ММО 4x4. По рисунку видно, что наименьшей вероятностью ошибки обладает Differential *SSK* (*DSSK*) модуляция. При неизменной вероятности ошибки *DSSK* имеет выигрыш ОСШ в 3 дБ относительно *SSK*-модуляции, *SSK* имеет выигрыш ОСШ в 2 дБ относительно *QSSK* (*Quadrature SM*) модуляции, *QSSK* имеет выигрыш ОСШ в 1 дБ относительно *GSSK* (*Generalized SM*) модуляции, *GSSK* имеет выигрыш ОСШ в 2 дБ относительно *GQSSK*-модуляции [4, 7].

Имитационное моделирование исследуемых систем производится в MATLAB на основе написанных авторами М-функций, по которым в дальнейшем строятся все нижеприведенные графики.

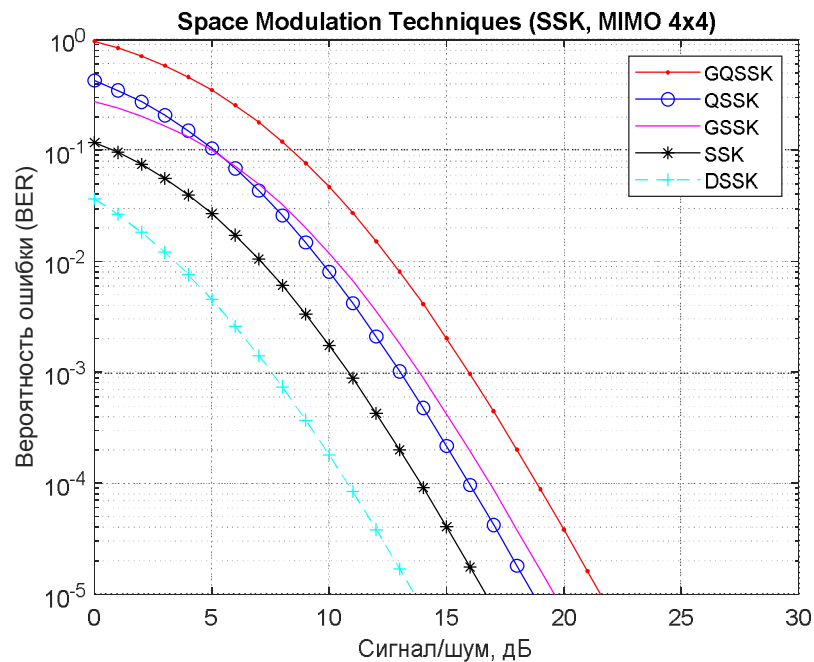


Рис. 6. Средняя вероятность ошибки для различных типов SSK-модуляции

В общем случае технологий SM совместно используются режимы обычной передачи сигнальных битов и антенных индексов. Здесь также имеются варианты, связанные с упорядоченным распределением пассивных и активных антенн [3, 4], детальное рассмотрение которых выходит за рамки статьи. Ниже для примера приводится таблица распределения символьных и пространственных битов при передаче технологией SM-сигналов модуляцией 4-QAM антенной решеткой, состоящей из четырех излучающих элементов [5].

Таблица 1. Распределение символьных и пространственных битов для $N_t = 4$ и 4-QAM-модуляции

Информационные биты	Символьные биты	Символы	Пространственные биты	Антенные индексы
0000	00	$+1+j$	00	1
0001	01	$-1+j$	00	1
0010	10	$-1-j$	00	1
0011	11	$+1-j$	00	1
0100	00	$+1+j$	01	2
0101	01	$-1+j$	01	2
0110	10	$-1-j$	01	2
0111	11	$+1-j$	01	2
1000	00	$+1+j$	10	3
1001	01	$-1+j$	10	3
1010	10	$-1-j$	10	3
1011	11	$+1-j$	10	3
1100	00	$+1+j$	11	4
1101	01	$-1+j$	11	4
1110	10	$-1-j$	11	4
1111	11	$+1-j$	11	4

Как следует из таблицы, передача информационных бит производится в определенном порядке: сначала передаются пространственные биты с антенным индексом, затем символьные биты звездной диаграммы синхронно с пространственными битами.

При оценке ошибок и емкости канала ММО при SM модель канала отличается от традиционной модели (1) уточнением:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{H}_l \mathbf{S}_l + \mathbf{n}. \quad (6)$$

$\mathbf{S}_l = s_l$ – комплексные сигнальные символы звездной диаграммы (QAM, PSK и т.д.), которые передаются активными антеннами с индексом l по нормализованному каналу $\mathbf{H}_l = h_l$, равному пространственному символу. При этом $\mathbf{n}$ – матрица аддитивного белого гауссовского шума с нулевым средним, ковариационной матрицей $\sigma_n^2 I_{N_r}$ и $OCSI = 1/\sigma_n^2$.

Вероятность ошибки при приеме сигналов SMTs и их передаче через плоский канал Релея в общем виде определяется уравнением [4]:

$$P_{error} = Q\left(\sqrt{\frac{\|h_\ell s_l - h_{\tilde{\ell}} s_{\tilde{l}}\|_F^2}{2\sigma_n^2}}\right) = Q\left(\sqrt{\sum_{r=1}^{N_r} \gamma_r^{SM}}\right) = Q\left(\sqrt{v^{SM}}\right), \quad (7)$$

где h_l – l -ый столбец матрицы $\mathbf{H}$ размера $N_r \times N_t$, пространственный символ передачи, $h_{\tilde{\ell}}$ – $\tilde{l}$ -ый столбец матрицы $\mathbf{H}$ размера $N_r \times N_t$, пространственный символ приема,

$$Q(x) = \frac{1}{2\pi} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt,$$

$$\gamma_r^{SM} = \frac{|h_{\ell,r} s_l - h_{\tilde{\ell},r} s_{\tilde{l}}|^2}{2\sigma_n^2},$$

где $h_{\ell,r}$ – r -ый элемент вектор-столбца $\mathbf{h}_\ell$ длины N_r ,

а $v^{SM} = \sum_{r=1}^{N_r} \gamma_r^{SM}$ – случайная величина, распределенная по критерию «хи-квадрат», с функцией распределения вероятностей

$$p_{v^{SM}}(v^{SM}) = \frac{1}{\Gamma(N_r)(\bar{\gamma}^{SM})^{N_r}} (v^{SM})^{(N_r-1)} \exp\left(-\frac{v^{SM}}{\bar{\gamma}^{SM}}\right),$$

где $\bar{\gamma}^{SM}$ – средняя величина, заданная выражением

$$\bar{\gamma}^{SM} = \frac{1}{2\sigma_n^2} \times \begin{cases} |s_l - s_{\tilde{l}}|^2 & \text{при } \tilde{l} = l, \\ |s_l|^2 + |s_{\tilde{l}}|^2 & \text{при } \tilde{l} \neq l. \end{cases}$$

Таким образом, средняя вероятность ошибки будет равна

$$P((\mathbf{h}_\ell, s_l) \rightarrow (\mathbf{h}_{\tilde{\ell}}, s_{\tilde{l}})) = E_{v^{SM}} \{P_{error}\} = (\alpha_a^{SM})^{N_r} \sum_{i=0}^{N_r-1} \binom{N_r-1+i}{i} [1 - \alpha_a^{SM}]^i, \quad (8)$$

где $\alpha_a^{SM} = \frac{1}{2} \left(1 - \sqrt{\frac{\tilde{\gamma}^{SM}/2}{1 + \tilde{\gamma}^{SM}/2}}\right)$. Отметим, что если v – случайная величина, распределенная по критерию «хи-квадрат», тогда

$$\begin{aligned}
E_v \{Q(\sqrt{v})\} &= \int_v p_v(v) Q(\sqrt{v}) dv = \\
&= \int_v \frac{1}{\Gamma(N_r)(\tilde{\gamma})^{N_r}} v^{(N_r-1)} \exp\left(-\frac{v}{\tilde{\gamma}}\right) Q(\sqrt{v}) dv = \\
&= (\alpha_a)^{N_r} \sum_{i=0}^{N_r-1} \binom{N_r-1+i}{i} [1-\alpha_a]^i,
\end{aligned} \tag{9}$$

$$\text{где } \alpha_a = \frac{1}{2} \left(1 - \sqrt{\frac{\tilde{\gamma}/2}{1+\tilde{\gamma}/2}} \right).$$

Расчет по формулам (7) – (9) при различном числе антенн приведен на рис. 7.

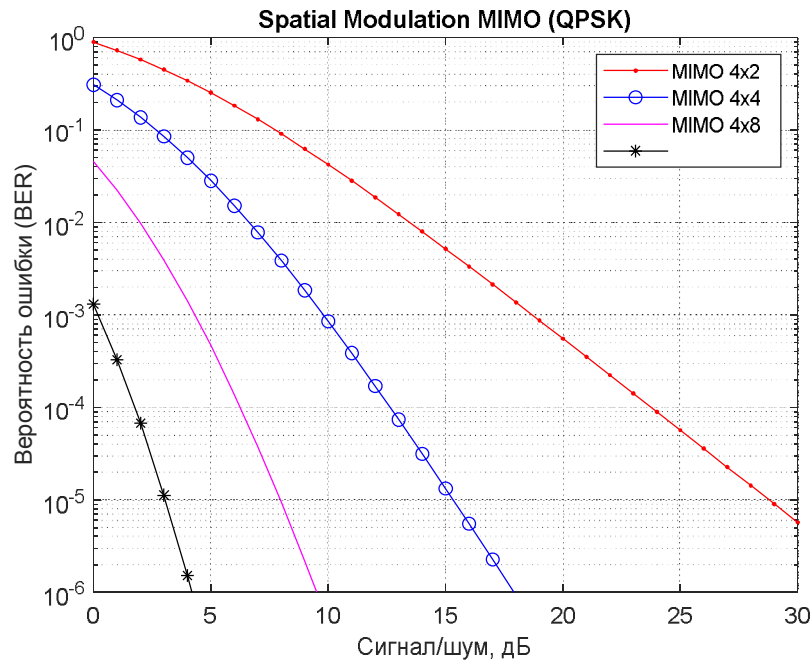


Рис. 7. Средняя вероятность ошибки для SM при различном числе приемных антенн N_r

И наконец, средняя вероятность ошибки для SM-модуляции при передаче сигналов через канал Релея с требуемой спектральной эффективностью определяется уравнением:

$$BER = \frac{1}{2^\eta} \sum_{\ell, l} \sum_{\ell, l} \frac{e_{\ell, l}^{\ell, l}}{\eta} (\alpha_a^{SM})^{N_r} \sum_{i=0}^{N_r-1} \binom{N_r-1+i}{i} [1-\alpha_a^{SM}]^i, \tag{10}$$

где η – требуемая спектральная эффективность для SM-модуляции,

$e_{\ell, l}^{\ell, l}$ – количество ошибочных бит, связанных с событием вероятности ошибки.

Уравнение (10) и рис. 7 показывают, что с увеличением числа приемных антенн существенно уменьшается вероятность ошибки. Это справедливо для всех вариантов SM модуляции, поэтому далее ограничимся сравнительными оценками вероятности ошибки для различных вариантов SM для некоррелированного канала Релея при полной информации на передатчике и приемнике.

На рис. 8 приведено сравнение большинства типов пространственной модуляции для системы MIMO 4x4 и используемой модуляции QPSK/4-QAM. По рисунку видно, что наименьшей вероятностью ошибки обладает DSM-модуляция. При неизменной вероятности ошибки DSM имеет выигрыш ОСШ в 3 дБ относительно SM-модуляции, SM имеет выигрыш ОСШ в 1 дБ относительно GSM-модуляции, GSM имеет выигрыш ОСШ в 2 дБ относительно QSM-модуляции, QSM имеет выигрыш ОСШ в 2 дБ относительно GQSM-модуляции.

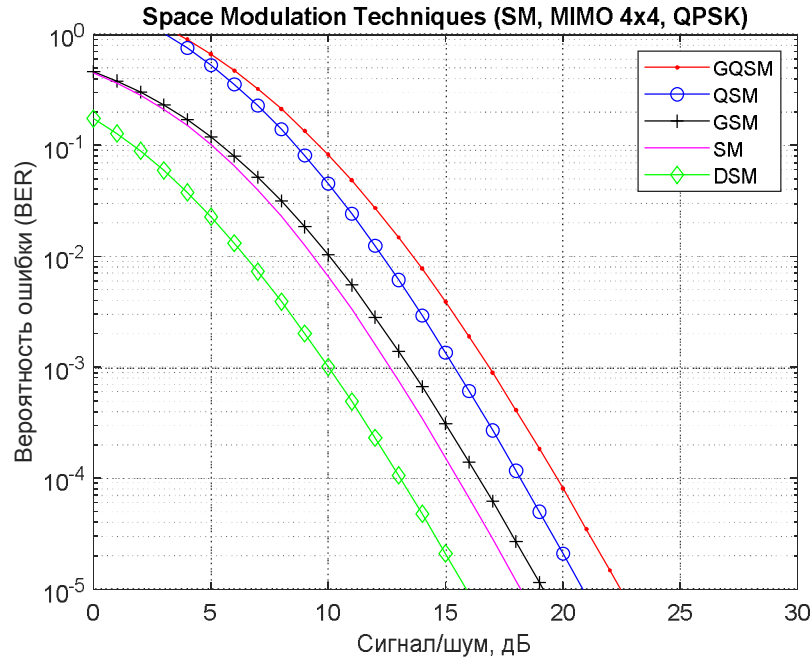


Рис. 8. Средняя вероятность ошибки для различных типов SM-модуляции

Немаловажную роль в современных технологиях системы MIMO играет анализ производительности сети. Для понимания разницы между классической системой MIMO и SMTs получена взаимная информация о пространственном мультиплексировании. В системах SMX не существует пространственных символов, а только символы сигналов. Входящая последовательность бит данных модулируется комплексными символами сигналов. И эти символы используются только для передачи информации: N_t символов будет одновременно передано с N_t передающих антенн.

По определению взаимная информация $I(\mathbf{X}; \mathbf{Y})$ – это количество информации, полученной о переданном векторном пространстве $\mathbf{X}$, зная принятое векторное пространство $\mathbf{Y}$. Взаимная информация для SMX-систем определяется формулой:

$$I(\mathbf{X}; \mathbf{Y}) = E_{\mathbf{H}} \{I(\mathbf{X}; \mathbf{Y} | \mathbf{H})\} = E_{\mathbf{H}} \{H(\mathbf{Y} | \mathbf{H}) - H(\mathbf{Y} | \mathbf{X}, \mathbf{H})\}, \quad (11)$$

где $H(\cdot)$ – функция энтропии.

В SMTs информационные биты модулируются пространственным символом $\mathcal{H}_\ell$ и символом сигнала $\mathcal{S}_\ell$. Поэтому в этом случае взаимная информация – это количество информации, полученной о пространственной и сигнальной диаграмме созвездий $\mathcal{H}$ и $\mathcal{S}$, зная принятое векторное пространство $\mathbf{Y}$. Тогда взаимная информация определяется формулой:

$$I(\mathcal{H}, \mathcal{S}; \mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y} | \mathcal{H}, \mathcal{S}). \quad (12)$$

При вычислении емкости сети SMTs информационные биты преобразуются в зависимости от индекса активной передающей антенны ℓ , а не от разных $\mathcal{H}_\ell$ пространственных символов. Следовательно, взаимная информация будет записана как

$$I(\ell, \mathcal{S}; \mathbf{Y} | \mathbf{H}) = E_{\ell} \{I(\mathcal{S}, \mathbf{Y} | \ell, \mathbf{H})\} + I(\ell; \mathbf{Y} | \mathbf{H}). \quad (13)$$

Емкость системы SMTs получается максимизацией взаимной информации с выбором функции распределения вероятности $p_{\mathcal{S}}$ и будет равна:

$$C = E_{\mathbf{H}} \left\{ \max_{p_{\mathcal{S}}} \{I(\ell, \mathcal{S}; \mathbf{Y} | \mathbf{H})\} \right\} = E_{\mathbf{H}} \left\{ \max_{p_{\mathcal{S}}} \{E_{\ell} \{I(\mathcal{S}, \mathbf{Y} | \ell, \mathbf{H})\}\} + I(\ell; \mathbf{Y} | \mathbf{H}) \right\} = E_{\mathbf{H}} \{C_1 + C_2\}, \quad (14)$$

где предполагается, что функция распределения вероятности p_s , которая максимизирует C_1 , также будет максимизировать и C_2 .

На рис. 9 отображена емкость системы SMTs для различных типов SSK (а) и SM (б) модуляций для MIMO 4x4 и 4-QAM/QPSK схемы модуляции.

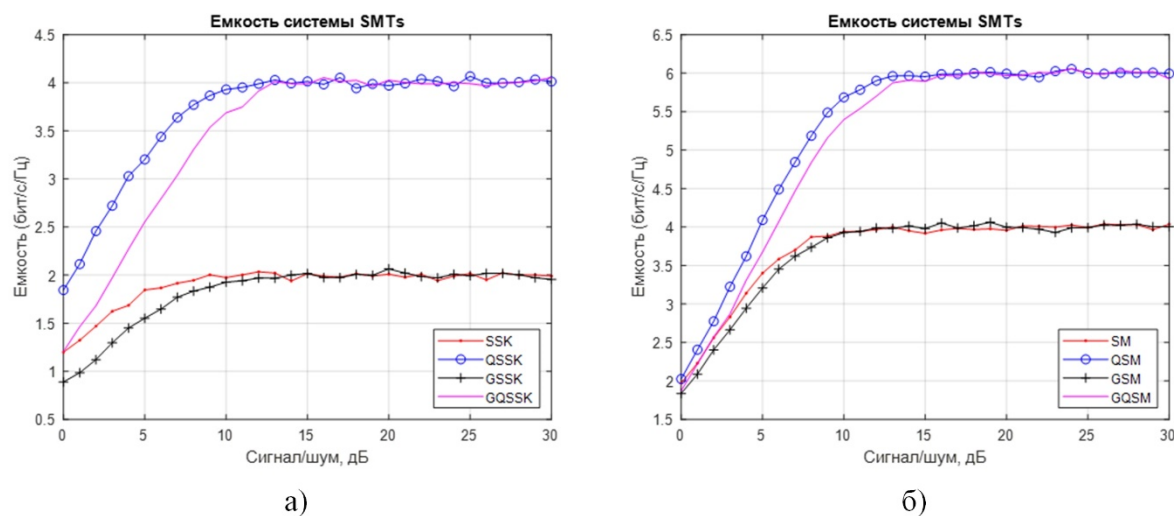


Рис. 9. Емкость системы SMTs

Для сравнения с SMTs на рис. 10 приводится емкость М-арных схем модуляций. Видно, что емкость SSK- и GSSK-систем соответствует емкости для схемы модуляции QPSK, а емкость QSSK и GQSSK в 2 раза выше. Для систем же SM и GSM емкость выше в 2 раза относительно QPSK, а для QSM и GQSM – в 3 раза.

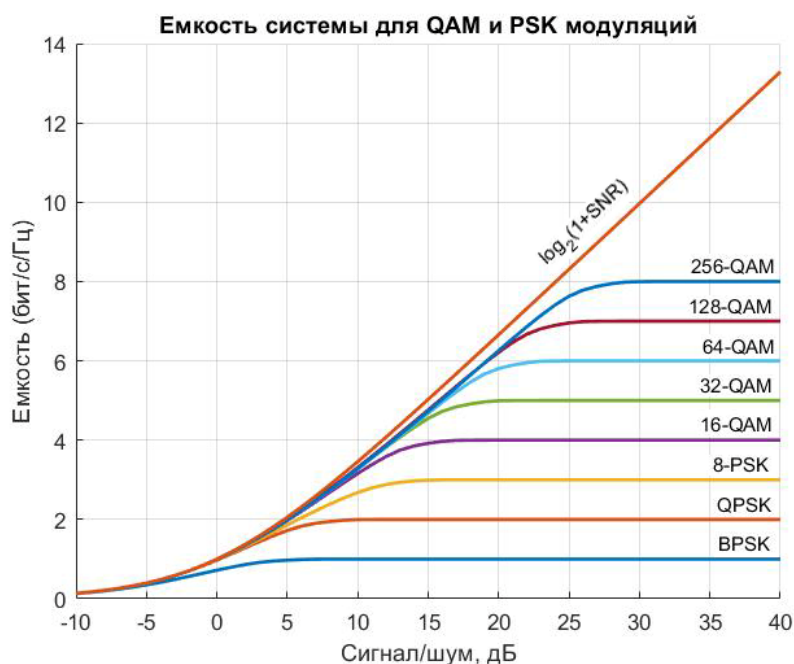


Рис. 10. Емкость системы для М-арных схем модуляций

На рис. 11 отображена емкость системы MIMO в зависимости от используемого числа приемных и передающих антенн.

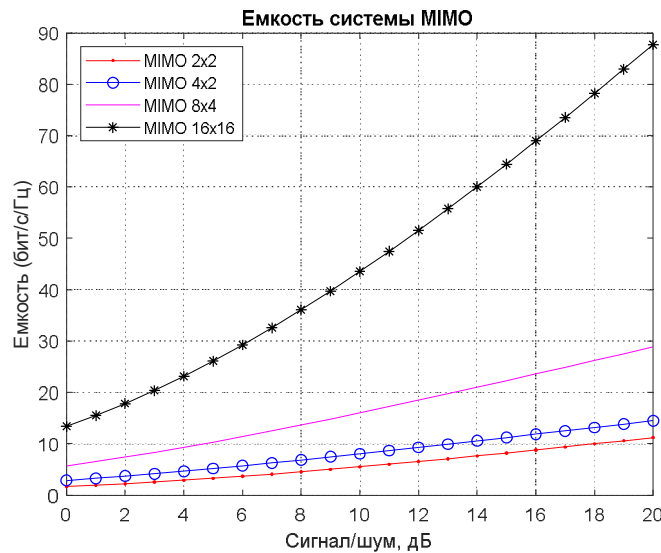


Рис. 11. Емкость системы MIMO

5. Характеристики технологии селекции передающих антенн

Метод *селекции передающих антенн* (*TAS – Transmit Antenna Selection*) заключается в том, что антенная система представляется в виде выбранной последовательности одиночных излучателей информационных сигналов. Выбор излучателя осуществляется на основе параметров канала, например, «хорошего канала», который определяется приемной антенной решеткой и по обратному каналу передается передатчику для выбора активной антенны. Этот метод напоминает схему SIMO (одна передающая антенна и много приемных) с *MRC* селективной обработкой в приемнике, но с существенным дополнением – обработка сигнала проводится не по одной искаженной копии, а по нескольким, что уменьшает вероятность ошибки.

В основу моделирования производительности сети положен неселективный канал Релея с комбинаторной обработкой узкополосных сигналов в приемнике.

Сравнение методов формирования сигнала на передачу приведено на рис. 12, что позволяет выбрать наилучший из методов, получая наименьшую вероятность ошибки.

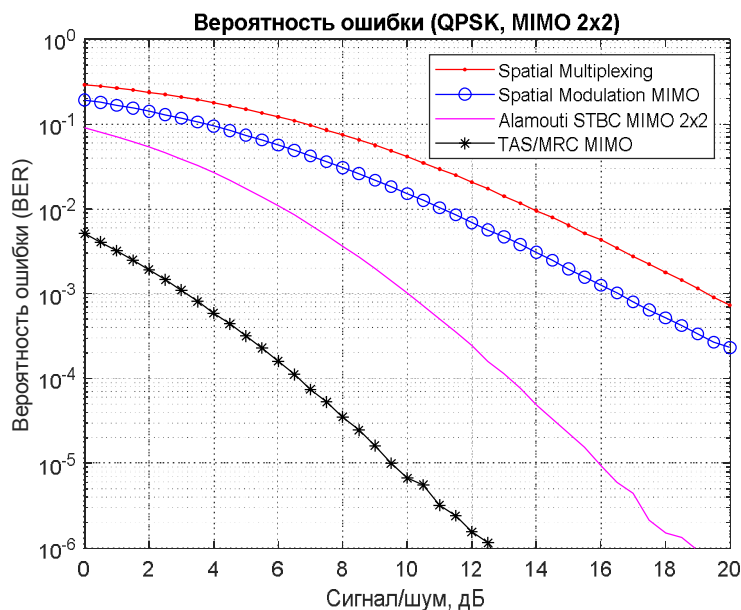


Рис. 12. Вероятность ошибки для различных систем MIMO

6. Заключение

В статье представлены результаты сравнительного анализа новых технологий ММО – пространственной модуляции и селекции передающих антенн с действующими технологиями – пространственным мультиплексированием и пространственно-временным кодированием. Полученные расчетные графики показывают, что технологии SM имеют преимущества в емкости системы и энергетической эффективности, а технологии TAS – в меньшей вероятности ошибки. Эти преимущества возрастают с увеличением числа антенн, что гармонирует с требованиями пятого поколения.

Литература

1. *Kumbhani B., Kshetrimayum R. S.* MIMO Wireless Communications over Generalized Fading Channels. CRC Press, 2017.
2. *Safak M.* Digital Communications. Wiley, 2017.
3. *Wen M., Cheng X., Yang L.* Index Modulation for 5G Wireless Communications. Springer, 2017.
4. *Mesleh R., Alhasssi A.* Space Modulation Techniques. Wiley, 2018.
5. *Kanatas A. G., Nikita K. S., Mathiopoulos P.* New Directions in Wireless Communications Systems. From Mobile to 5G. CRC Press, 2018.
6. *Duong T. Q., Chu X., Suraveera H. A.* Ultra-dense Networks for 5G and Beyond. Wiley, 2019.
7. *Петров В. П., Якушев И. Ю.* Исследование производительности ММО систем пятого поколения в неселективном канале Релея // VI Международная научно-практическая конференция «Технологии разработки и отладки сложных технических систем» на базе МГТУ им. Н.Э. Баумана, 2019.
8. *Гантмахер Ф. Р.* Теория матриц. М.: Наука, 1967.

Статья поступила в редакцию 22.05.2019.

Петров Виктор Петрович

д.т.н., профессор, профессор кафедры систем мобильной связи СибГУТИ (630102, Новосибирск, ул. Кирова, 86), e-mail: petrov@sibsutis.ru.

Якушев Игорь Юрьевич

аспирант кафедры систем мобильной связи СибГУТИ, e-mail: yakushevigor@mail.ru.

Modern Technologies in MIMO System

V.P. Petrov, I.Yu. Yakushev

In the paper MIMO systems with various technologies for optimizing system performance in a non-selective Rayleigh channel have been considered. The focus is on a relatively new antenna technologies recommended for fifth-generation networks: Spatial Modulation and Transmit Antenna Selection technologies. Formulas of BER and channel capacity have been given as well as graphs based on them for characteristics comparison.

Keywords: 5G, BER, MIMO, Rayleigh fading channel, channel capacity, antenna technology.