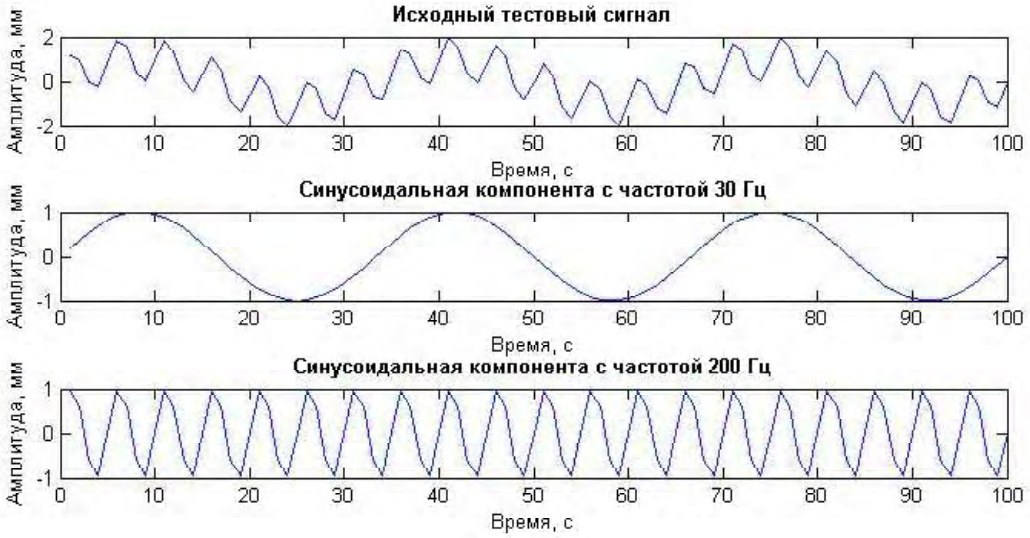


Рис. 1. Пример тестового сигнала

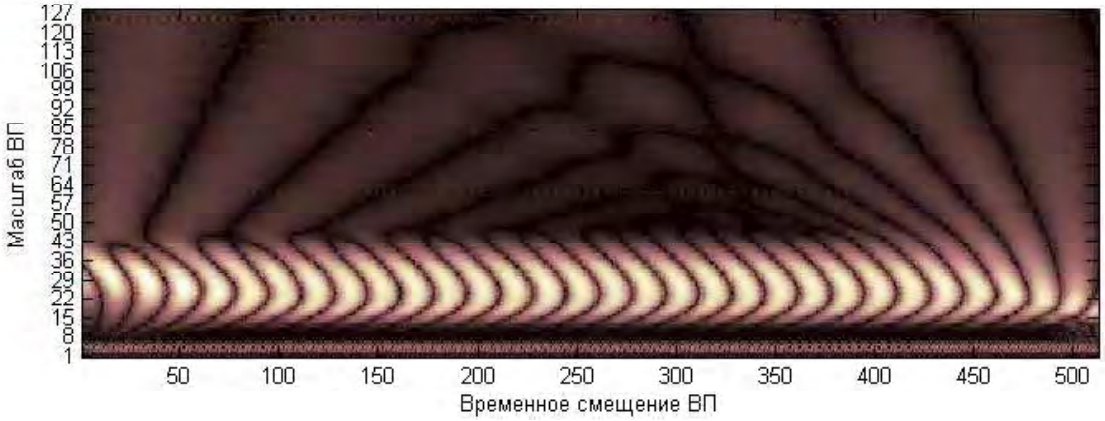


Рис. 2. Вейвлет-спектр тестового сигнала

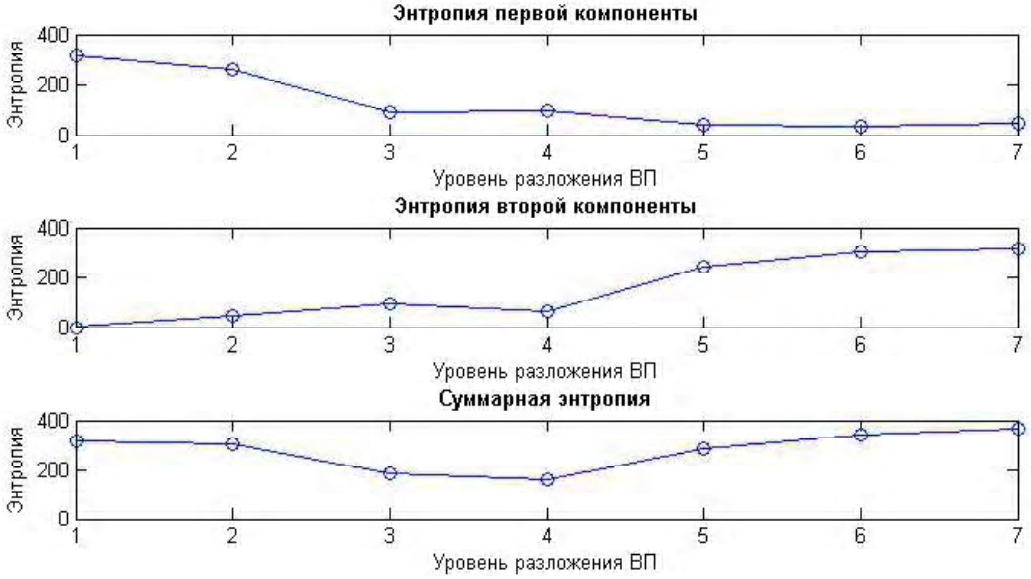


Рис. 3. Вычисление энтропии Шеннона уровней вейвлет-разложения сигнала

Рассмотрим изменение энтропии компонент сигнала. Значение энтропии первой компоненты уменьшается, а второй компоненты увеличивается. Изменение энтропии происходит более резко, чем энергии. Но суммарная энтропия двух компонент постепенно уменьшается до определенного уровня, а потом происходит увеличение значений. Можно предположить, что «переломная точка» отображает наиболее неустойчивое состояние системы, в котором возможно выделение компонент (рис. 5). Согласно теории К. Шеннона, прирост информации равен утраченной неопределённости системы (2), поэтому на этом уровне декомпозиции остановился прирост информации. Система перешла в абсолютно неустойчивое состояние. Из вышеперечисленного можно сделать следующий вывод: выделение компонент сигнала можно выполнить на уровне декомпозиции с минимальным значением суммарной энтропии обоих компонент.

5 ПРОЦЕДУРА ВЫДЕЛЕНИЯ КОМПОНЕНТ СИГНАЛА

Процедура выделения компонент сигнала состоит из следующих этапов:

1. Установить уровень разложения $i=1$.

2. Проверка: $i \le 2^{N/2}$. Если выполняется условие – переход к п. 3, иначе – п. 11.

3. Выполнить вейвлет-преобразование сигнала на i-уровне.

4. Выделить первую компоненту сигнала – оставить без изменений коэфициенты аппроксимации и обнулить коэффициенты детализации.

5. Выделить вторую компоненту сигнала – вычесть из общего сигнала первую компоненту.

6. Вычислить энтропию первой компоненты.

7. Вычислить энтропию второй компоненты.

8. Вычислить суммарную энтропию обоих компонент.

9. Проверка: если суммарная энтропия на i-уровне меньше суммарной энтропии на $(i–1)$-уровне – перейти к п. 10, иначе к п. 11.

10. Увеличить уровень вейвлет-разложения $i=i+1$. Перейти к п. 2.

11. Установить оптимальный уровень выделения компонент в $(i–1)$-уровень.

12. Останов.

ВЫВОДЫ

В результате проведения исследовательской работы получены следующие выводы:

– исследовано поведение энергии и энтропии сигнала на этапах его декомпозиции;

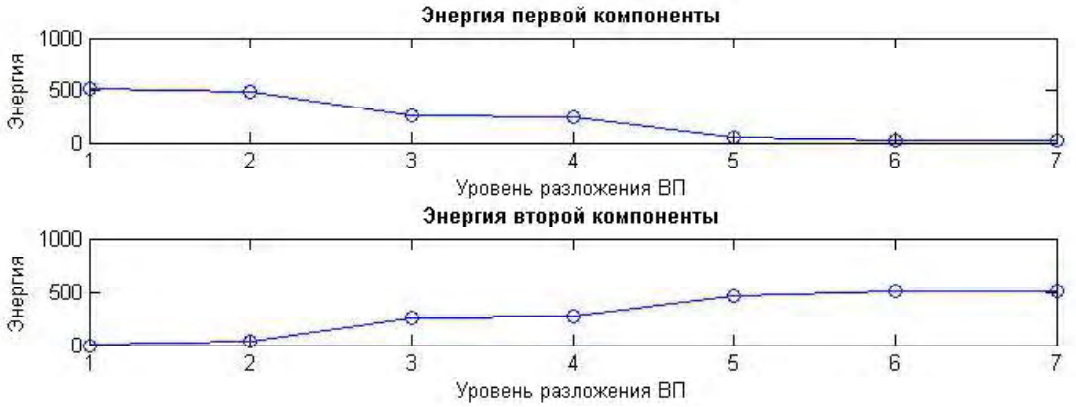


Рис. 4. Вычисление энергии уровней вейвлет-разложения сигнала

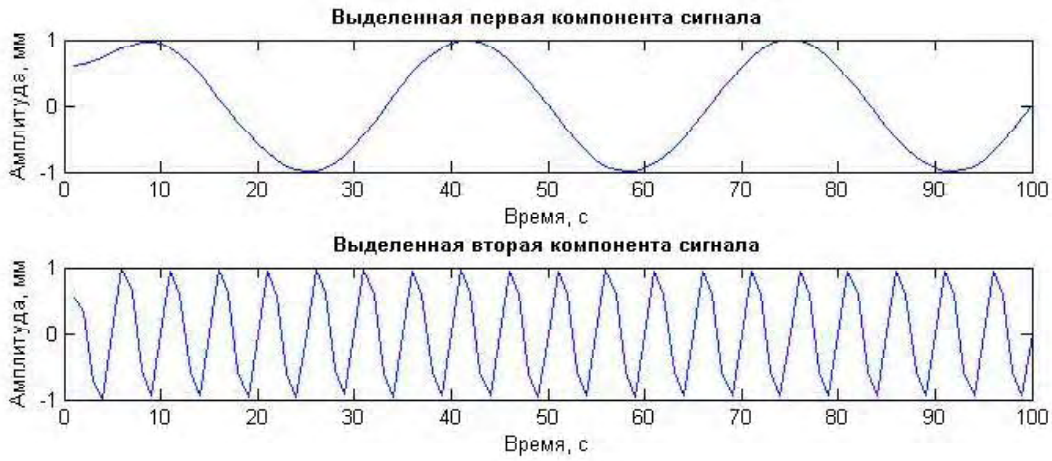


Рис. 5. Выделение компонент сигнала на 4 уровне вейвлет-разложения

– модифицирован метод Кофмана для задачи выделения компонент сигнала;

– впервые предложен метод разделения сигнала на компоненты на основе вейвлет-декомпозиции и теории информации;

– разработана процедура выделения компонент сигнала.

СПИСОК ЛИТЕРАТУРЫ

1. *Sen, M.* Real-time digital signal processing. Implementations and applications / M. Sen. – Wiley, 2006. – 667 p.
2. *Misiti, M.* Wavelet and their applications / M. Misiti, G. Oppenheim. – USA : ISTE, 2007. – 330 p.
3. *Percival, D.* Wavelet methods for time series analysis / D. Percival, A. Walden. – Cambridge university press, 2003. – 566 p.
4. *Mallat, S.* A wavelet tour of signal processing / S. Mallat. – USA: Academic Press, 1998 – 805 p.
5. *Coifman, R. R.* Entropy-based algorithms for best basis selection / R. R. Coifman, M. V. Wickerhauser // IEEE Trans. on Inf. Theory. – 1992. – Vol. 38 (2). – P. 713–718.
6. *Martin, N.* Mathematical theory of entropy / N. Martin. – Cambridge university press, 1984. – 258 p.
7. *Muller, I.* Entropy and energy / I. Muller, W. Weiss. – Springer Press, 2005. – 273 p.
8. *Fang, S.* Entropy optimization and mathematical programming / S. Fang, J. Rajasekera. – USA: Kluwer academic publishers, 1997. – 273 p.
9. *Shannon, K.* A Mathematical theory of communication / K. Shannon // The Bell System Technical Journal. – 1948. – Vol. 27. – P. 379–423.

Стаття надійшла до редакції 25.10.2013.

Дубровін В. І.[1], Твердохліб Ю. В.[2]

[1]Канд. техн. наук, професор, Запорізький національний технічний університет, Україна

[2]Аспірант, Запорізький національний технічний університет, Україна

ДОСЛІДЖЕННЯ ЗМІН ЕНТРОПІЇ ТА ЕНЕРГІЇ НА ЕТАПАХ ДЕКОМПОЗИЦІЇ СИГНАЛУ

Проведено дослідження виявлення змін ентропії та енергії сигналу на етапах його декомпозиції. Запропоновано метод ділення сигналу складної форми на незалежні складові за допомогою математичного апарату вейвлет-перетворення та теорії інформації.

Ключові слова: вейвлет-перетворення, розділення сигналу, оптимальна декомпозиція сигналу.

Dubrovin V. I.[1], Tverdohleb J. V.[2]

[1]Ph.D. in Engineering, professor, Zaporozhye national technical university, Ukraine

[2]Aspirant, Zaporozhye national technical university, Ukraine

RESEARCH OF CHANGES OF ENTROPY AND ENERGY ON SIGNAL DECOMPOSITION

Research of changes of entropy and energy on signal decomposition is presented. Coifman's method for the purpose for signal delineation was modified. The method of complex signal delineation based on wavelet transformation and information theory is proposed. The algorithm for the signal components separation is developed. The proposed method based on research of total entropy of both signal formed components. The first component is the reconstructed source signal after wavelet transformation on the current decomposition level in which the approximation coefficients must be equals to zero. The second component is the residue from the deduction of the source signal and the second component on the current decomposition level. The turning point of the total entropy curve denotes the most unstable system state in which the components can be identified. The most unstable system state is system state having zero increment of information. Entropy increments until the turning point and it decrease after. The sum of the sinusoids was used as test signal. The results of the method have high accuracy.

Keywords: wavelet transformation, signal delineation, optimal signal decomposition.

REFERENCES

1. Sen M. Real-time digital signal processing. Implementations and applications. Wiley, 2006, 667 p.
2. Misiti M., Oppenheim G. Wavelet and their applications. USA : ISTE, 2007, 330 p.
3. Percival D., Walden A. Wavelet methods for time series analysis. Cambridge university press, 2003, 566 p.
4. Mallat S. A wavelet tour of signal processing. USA: Academic Press, 1998, 805 p.
5. Coifman R. R., Wickerhauser M. V. Entropy-based algorithms for best basis selection, *IEEE Trans. on Inf. Theory*, 1992, Vol. 38 (2), pp. 713–718.
6. Martin N. Mathematical theory of entropy. Cambridge university press, 1984, 258 p.
7. Muller I., Weiss W. Entropy and energy, Springer Press, 2005, 273 p.
8. Fang S., Rajasekera J. Entropy optimization and mathematical programming. USA : Kluwer academic publishers, 1997, 273 p.
9. Shannon K. A Mathematical theory of communication, *The Bell System Technical Journal*, 1948, Vol. 27, pp. 379–423.

УДК 519.812.3

Колпакова Т. А.

Аспирант, Запорожский национальный технический университет, Украина, E-mail: tanya_kolpakova@mail.ru

ОПТИМИЗАЦИЯ РАСПРЕДЕЛЕНИЯ ОБЪЕМА ЗАКАЗА МЕЖДУ НЕСКОЛЬКИМИ ПОСТАВЩИКАМИ В ТЕНДЕРЕ

Рассматривается проблема оптимального распределения объема заказа между несколькими поставщиками, участвующими в тендере. Предложен метод многокритериальной оптимизации распределения объема заказа между несколькими поставщиками в тендере, который учитывает относительные коэффициенты важности критериев, по которым производится оптимизация.

Ключевые слова: многокритериальная оптимизация, распределение объема заказа, критерий оптимальности, относительные приоритеты критериев.

ВВЕДЕНИЕ

После проведения тендера в процессе уточнения условий заключения контракта нередко появляются новые неприемлемые для подрядчика условия, из-за чего приходится обращаться к следующему по рейтингу участнику либо проводить новый тендер. Чтобы избежать подобных ситуаций организаторы тендеров все чаще предпочитают заключать контракты не с одним, а с несколькими поставщиками.

Еще на этапе проведения тендера можно предусмотреть, как распределить весь необходимый объем заказа между некоторым числом подрядчиков [1]. Однако, это возможно только в том случае, если участники предлагают одинаковые товары или услуги.

При решении данной задачи чаще всего проводится оптимизация по двум критериям, которые можно однозначно представить в количественном виде: по денежным затратам и по времени, необходимому для выполнения заказа. Важно то, что эти критерии противоречивы, и оптимизация по каждому из них приводит к ухудшению значений по другому критерию и соответственно к различным решениям задачи распределения [2].

1 ПОСТАНОВКА ЗАДАЧИ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ РАСПРЕДЕЛЕНИЯ ОБЪЕМА ЗАКАЗА МЕЖДУ НЕСКОЛЬКИМИ ПОСТАВЩИКАМИ В ТЕНДЕРЕ

Если заказчик желает заключить контракт не с одним, а сразу с i участниками, то задачу многокритериальной оптимизации распределения объема всего заказа U между несколькими поставщиками можно представить в виде задачи минимизации суммы всех затрат P частей заказа u_i, а также минимизации времени Δt_i, необходимого для выполнения заказа. При этом следует учитывать относительные приоритеты критериев стоимости и времени исполнения заказа.

В данном случае, если участники будут выполнять поставку товаров или услуг одновременно, то общее время выполнения полного объема заказа будет равно максимальному из периодов времени, заявленных заказчиками.

$$\min_{\Omega} \Phi, \tag{1}$$

$$\Phi = \{\Phi_k \mid k = 1..2\}, \tag{2}$$

$$\Phi_1 = \sum_i u_i P_i, \tag{3}$$

$$\Phi_2 = \max_i \Delta t_i, \tag{4}$$

$$\Omega : U = \sum_i u_i, \tag{5}$$

где $\Phi = (\Phi_1, \Phi_2)$ – векторный критерий оптимальности, Φ_k – частные критерии оптимальности.

Поиск оптимального распределения является итерационным процессом перебора на некотором дискретном множестве вариантов решений. При поиске оптимального решения показателем положительного результата итерации является то, что относительный уровень снижения качества по одному частному критерию не превосходит относительного уровня повышения качества по другому частному критерию.

Поскольку в процессе проведения тендера определяется, что одни критерии важнее других, следует определить относительные приоритеты критериев стоимости и времени исполнения заказа d_k, причем

$$\sum_k d_k = 1. \tag{6}$$

Пусть существуют решения Y_a и Y_b – способы распределения объема заказа U на части u_i. Мера относительного изменения (снижения или повышения) качества решения при переходе от решения Y_a к Y_b по каждому из критериев составляет:

$$\Delta\tilde{\Phi}_k(Y_a, Y_b) = \frac{\Delta\Phi_k(Y_a, Y_b)}{\left| \max_{Y \in \{Y_a, Y_b\}} \Phi_k \right|}, \tag{7}$$

где $\Delta\Phi_k(Y_a, Y_b) = \Phi_k(Y_a) - \Phi_k(Y_b)$ – абсолютные изменения значений частных критериев оптимальности при переходе от решения Y_a к решению Y_b.

Максимальное снижение качества решения при переходе от решения Y_a к решению Y_b определяется как:

$$\Delta\tilde{\Phi}_{\min}(Y_a, Y_b) = \min_k d_k \cdot \Delta\tilde{\Phi}_k(Y_a, Y_b). \quad (8)$$

Аналогично максимальное повышение качества решения при переходе от решения Y_a к решению Y_b составляет:

$$\Delta\tilde{\Phi}_{\max}(Y_a, Y_b) = \max_k d_k \cdot \Delta\tilde{\Phi}_k(Y_a, Y_b). \quad (9)$$

Решение Y_b превосходит решение Y_a ($Y_b \succ Y_a$), если:

$$\left|\Delta\tilde{\Phi}_{\max}(Y_a, Y_b)\right| > \left|\Delta\tilde{\Phi}_{\min}(Y_a, Y_b)\right|. \quad (10)$$

И, соответственно, решение Y_a превосходит решение Y_b ($Y_a \succ Y_b$), если:

$$\left|\Delta\tilde{\Phi}_{\max}(Y_a, Y_b)\right| \le \left|\Delta\tilde{\Phi}_{\min}(Y_a, Y_b)\right|. \quad (11)$$

3 МЕТОД МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ РАСПРЕДЕЛЕНИЯ ОБЪЕМА ЗАКАЗА МЕЖДУ НЕСКОЛЬКИМИ ПОСТАВЩИКАМИ В ТЕНДЕРЕ

Выбор начального решения может быть произведен случайным образом, однако эффективнее для начала поиска выбрать такое решение, при котором весь объем закупается у лучшего по предварительным оценкам поставщика.

Однако следует также учесть, что поставщик может ограничить минимальный объем, который он согласен поставить в рамках одного контракта. Значение части объема заказа u_i для каждого поставщика следует обговорить дополнительно еще на этапе проведения тендера.

Основные этапы предложенного метода:

Шаг 1. На первой итерации $r=1$ выбирается начальное решение задачи распределения Y_0.

Шаг 2. Вычисляются значения частных критериев оптимальности Φ_k для выбранного решения Y_0.

Шаг 3. Некоторым способом выбирается решение задачи распределения Y_r.

Шаг 4. Вычисляются значения частных критериев оптимальности Φ_k для решения Y_r.

Шаг 5. Согласно выражениям 7–11 определяется превосходящее решение Y_t.

Шаг 6. Если условие окончания поиска выполнено, то вычисления следует прекратить, а решение Y_t принять за приближенное решение задачи. В противном случае, следует положить $Y_0 = Y_t$ и перейти к шагу 3.

В качестве условия окончания итераций может быть использовано достижение предварительно заданного количества итераций, либо окончание полного перебора дискретного множества возможных решений Y_s.

Наилучшее решение Y_t будет содержать оптимальное распределение объема заказа U на i частей u_i.

3 РЕШЕНИЕ ЗАДАЧИ ЗАКАЗА КАБЕЛЕЙ ДЛЯ СТРОИТЕЛЬСТВА ПРОМЫШЛЕННОГО ОБЪЕКТА

Необходимо распределить объем заказа силовых кабелей для строительства крупного промышленного объекта. Тендерное задание включает следующую позицию:

Кабель силовой, 1 кВ, с медными жилами 2х185 мм. кв., в ПВХ не распространяющем горение (ВВГнг-1) – **1137 м.**

После проведения предварительного отбора, заказчики остановились на двух участниках:

1. «Энергопром» (Днепропетровск):
– цена за 1 ед. с НДС – **306,660 грн.;**
– общая сумма с НДС – **348 672,42 грн.;**
– срок исполнения – **30 дней.**
2. ООО «КОРТ» (Задачин):
– цена за 1 ед. с НДС – **344,66 грн.;**
– общая сумма с НДС – **391 878,97 грн.;**
– срок исполнения – **21 день.**

При этом коэффициенты относительных приоритетов критериев составляют:

– приоритет критерия «Стоимость заказа» $d_P = 0{,}54$;
– приоритет критерия «Срок исполнения» $d_t = 0{,}46$.

В таблицах 1 и 2 приведено множество решений распределения объема заказа с шагом дискретизации 0,2 и с учетом того, что поставщики не ограничили объем одной поставки.

Процесс поиска оптимального решения состоит из следующих шагов:

Шаг 1. За начальное решение Y_0 принимается распределение 1, при котором весь заказ выполняется вторым подрядчиком.

Шаг 2. Для Y_0 значения частных критериев оптимальности составляют:

$$\Phi_P = 391878{,}97 \text{ грн.},$$

$$\Phi_t = 21 \text{ день}.$$

Шаг 3. Выбирается значение Y_1 как распределение 2.

Шаг 4. Для Y_1 значения частных критериев оптимальности составляют:

$$\Phi_P = 0{,}2*348672{,}42 \text{ грн} + 0{,}8*391878{,}97 \text{ грн} = 383237{,}66 \text{ грн.},$$

$$\Phi_t = \max(\lceil 0{,}2*30 \rceil, \lceil 0{,}8*21 \rceil) = \max(6{,}17) = 17 \text{ дней}.$$

Шаг 5. Определяются меры изменения качества решения при переходе от Y_0 к Y_1.

$$\Delta\tilde{\Phi}_P(Y_0, Y_1) = \frac{391878{,}97 - 383237{,}66}{391878{,}97} = 0{,}0221,$$

$$\Delta\tilde{\Phi}_t(Y_0, Y_1) = \frac{21-17}{21} = 0{,}1905.$$

Таблица 1. Стоимость заказа из расчета его распределения на доли u_1 и u_2

№	Доля заказа u_1	Доля заказа u_2	Стоимость части заказа u_1	Стоимость части заказа u_2	Общая стоимость заказа
1	0	1	0,00	391 878,97	391 878,97
2	0,2	0,8	69 734,48	313 503,18	383 237,66
3	0,4	0,6	139 468,97	235 127,38	374 596,35
4	0,6	0,4	209 203,45	156 751,59	365 955,04
5	0,8	0,2	278 937,94	78 375,79	357 313,73
6	1	0	348 672,42	0,00	348 672,42

Таблица 2. Срок исполнения заказа из расчета его распределения на доли u_1 и u_2

№	Доля заказа u_1	Доля заказа u_2	Срок исполнения заказа u_1	Срок исполнения заказа u_2	Общий срок исполнения заказа
1	0	1	0	21	21
2	0,2	0,8	6	17	17
3	0,4	0,6	12	13	13
4	0,6	0,4	18	9	18
5	0,8	0,2	24	5	24
6	1	0	30	0	30

Шаг 6. Очевидно, что при переходе от Y_0 к Y_1 качество решения повышается по обоим параметрам, значит на данном этапе Y_1 – превосходящее решение, и дальнейший поиск следует продолжать относительно него.

Шаг 7. Выбирается значение Y_2 как распределение 3.

Шаг 8. Для Y_2 значения частных критериев оптимальности составляют:

$$\Phi_P = 0{,}4 * 348672{,}42 \text{ грн} + 0{,}6 * 391878{,}97 \text{ грн} = 374596{,}35 \text{ грн.},$$

$$\Phi_t = \max(\lceil 0{,}4 * 30 \rceil, \lceil 0{,}6 * 21 \rceil) = \max(12{,}13) = 13 \text{ дней.}$$

Шаг 9. Определяются меры изменения качества решения при переходе от Y_1 к Y_2.

$$\Delta\tilde{\Phi}_P(Y_1, Y_2) = \frac{3383237{,}66 - 374596{,}35}{383237{,}66} = 0{,}0225,$$

$$\Delta\tilde{\Phi}_t(Y_1, Y_2) = \frac{17 - 13}{17} = 0{,}2353.$$

Шаг 10. Очевидно, что при переходе от Y_1 к Y_2 качество решения повышается по обоим параметрам, значит на данном этапе Y_2 – превосходящее решение, и дальнейший поиск следует продолжать относительно него.

Шаг 11. Выбирается значение Y_3 как распределение 4.

Шаг 12. Для Y_3 значения частных критериев оптимальности составляют:

$$\Phi_P = 0{,}6 * 348672{,}42 \text{ грн} + 0{,}4 * 391878{,}97 \text{ грн} = 365955{,}04 \text{ грн.},$$

$$\Phi_t = \max(\lceil 0{,}6 * 30 \rceil, \lceil 0{,}4 * 21 \rceil) = \max(18{,}9) = 18 \text{ дней.}$$

Шаг 13. Определяются меры изменения качества решения при переходе от Y_2 к Y_3.

$$\Delta\tilde{\Phi}_P(Y_2, Y_3) = \frac{374596{,}35 - 365955{,}04}{374596{,}35} = 0{,}0231.$$

$$\Delta\tilde{\Phi}_t(Y_2, Y_3) = \frac{13 - 18}{18} = -0{,}2778;$$

Максимальное снижение качества решения при переходе от Y_2 к Y_3 составляет:

$$\Delta\tilde{\Phi}_{\min}(Y_2, Y_3) = 0{,}46 * (-0{,}2778) = -0{,}1278.$$

Максимальное повышение качества решения при переходе от Y_2 к Y_3 составляет:

$$\Delta\tilde{\Phi}_{\max}(Y_2, Y_3) = 0{,}54 * 0{,}0231 = 0{,}0125.$$

Поскольку

$$\left|\Delta\tilde{\Phi}_{\max}(Y_2, Y_3)\right| < \left|\Delta\tilde{\Phi}_{\min}(Y_2, Y_3)\right|,$$

превосходящим решением является Y_2.

Шаг 14. Выбирается значение Y_4 как распределение 5.

Шаг 15. Для Y_4 значения частных критериев оптимальности составляют:

$$\Phi_P = 0{,}8 * 348672{,}42 \text{ грн} + 0{,}2 * 391878{,}97 \text{ грн} = 357313{,}73 \text{ грн.},$$

$$\Phi_t = \max(\lceil 0{,}8 * 30 \rceil, \lceil 0{,}2 * 21 \rceil) = \max(24{,}5) = 24 \text{ дня.}$$

Шаг 16. Определяются меры изменения качества решения при переходе от Y_2 к Y_4.

$$\Delta\tilde{\Phi}_P(Y_2,Y_4)=\frac{374596{,}35-357313{,}73}{374596{,}35}=0{,}0461;$$

$$\Delta\tilde{\Phi}_t(Y_2,Y_4)=\frac{13-24}{24}=-0{,}4583.$$

Максимальное снижение качества решения при переходе от Y_2 к Y_4 составляет:

$$\Delta\tilde{\Phi}_{\min}(Y_2,Y_4)=0{,}46*(-0{,}45838)=-0{,}2108.$$

Максимальное повышение качества решения при переходе от Y_2 к Y_4 составляет:

$$\Delta\tilde{\Phi}_{\max}(Y_2,Y_4)=0{,}54*0{,}0461=0{,}0249.$$

Поскольку

$$\left|\Delta\tilde{\Phi}_{\max}(Y_2,Y_4)\right|<\left|\Delta\tilde{\Phi}_{\min}(Y_2,Y_4)\right|,$$

превосходящим решением является Y_2.

Шаг 17. Выбирается значение Y_5 как распределение 6.

Шаг 18. Для Y_5 значения частных критериев оптимальности составляют:

$$\Phi_P=348672{,}42 \text{ грн.},$$

$$\Phi_t=30 \text{ дней}.$$

Шаг 19. Определяются меры изменения качества решения при переходе от Y_2 к Y_5.

$$\Delta\tilde{\Phi}_P(Y_2,Y_5)=\frac{374596{,}35-348672{,}42}{374596{,}35}=0{,}0692,$$

$$\Delta\tilde{\Phi}_t(Y_2,Y_5)=\frac{13-30}{30}=-0{,}5667.$$

Максимальное снижение качества решения при переходе от Y_2 к Y_5 составляет:

$$\Delta\tilde{\Phi}_{\min}(Y_2,Y_5)=0{,}46*(-0{,}5667)=-0{,}2607.$$

Максимальное повышение качества решения при переходе от Y_2 к Y_5 составляет:

$$\Delta\tilde{\Phi}_{\max}(Y_2,Y_5)=0{,}54*0{,}0692=0{,}0374.$$

Поскольку

$$\left|\Delta\tilde{\Phi}_{\max}(Y_2,Y_5)\right|<\left|\Delta\tilde{\Phi}_{\min}(Y_2,Y_5)\right|,$$

превосходящим решением является Y_2.

Таким образом, за 19 шагов был совершен полный перебор дискретного множества решений, заданных в таблицах 1 и 2. После завершения перебора наилучшим решением оказалось решение Y_2, которое соответствует распределению 3. Оптимальным решением задачи будет заключение контракта на 40 % заказа (455 м кабеля) с подрядчиком «Энергопром», и на 60 % заказа (682 м кабеля) с подрядчиком ООО «КОРТ».

При этом общая стоимость заказа уменьшается на 4 %, и общий срок исполнения заказа уменьшается на 38 % относительно решения, при котором весь заказ выполняет подрядчик ООО «КОРТ».

ВЫВОДЫ

Предложенный метод многокритериальной оптимизации позволил распределить объем заказа между несколькими поставщиками согласно критериям стоимости и срока исполнения заказа. При этом он учитывает относительные коэффициенты важности критериев, что позволяет отыскать оптимальное решение в том случае, когда критерии не равнозначны для текущего тендера.

Использование предложенного метода позволяет еще на этапе проведения тендера минимизировать затраты на заказы, уменьшить общее время выполнения заказа относительно планового, и избежать возможных задержек в поставках.

СПИСОК ЛИТЕРАТУРЫ

1. *Guedas, B.* Compromise Based Evolutionary Multiobjective Optimization Algorithm for Multidisciplinary Optimization / Benoit Guedas, Xavier Gandibleux, Philippe Depince // Lecture Notes in Economics and Mathematical Systems. – 2011. – Volume 648. – P. 69–78.
2. *Seydel, J.* Multicriteria Support for Construction Bidding / J. Seydel, D. L. Olson // Mathematical and Computer Modelling. – 2001. – № 34. – P. 677–702.

Стаття надійшла до редакції 25.10.2013.

Колпакова Т. О.
Аспірант, Запорізького національного технічного університету, Україна

ОПТИМІЗАЦІЯ РОЗПОДІЛУ ОБСЯГУ ЗАМОВЛЕННЯ МІЖ ДЕКІЛЬКОХ ПОСТАЧАЛЬНИКІВ У ТЕНДЕРІ

Розглядається проблема оптимального розподілу обсягу замовлення між декількома постачальниками, які беруть участь у тендері. Запропоновано метод багатокритеріальної оптимізації розподілу обсягу замовлення між декількома постачальниками в тендері, який враховує відносні коефіцієнти важливості критеріїв, за якими проводиться оптимізація.

Ключові слова: багатокритеріальна оптимізація, розподіл обсягу замовлення, критерій оптимальності, відносні пріоритети критеріїв.

Kolpakova T. A.
Postgraduate student, Zaporizhzhya National Technical University, Ukraine

OPTIMIZATION OF DISTRIBUTION OF ORDERS AMONG MULTIPLE CONTRACTORS IN TENDERING

The problem of optimal distribution of orders among multiple contractors participating in the tender is considered.

During tendering when contract is prepared for signing, some additional requirements can appear. If the selected contractor can't satisfy them, client should choose another contractor or organize one more tender. To avoid this situation client may want to split job among several contractors.

A method for multi-objective optimization of distribution of orders among several contractors in the tender, which takes into account the relative importance of factors of optimization criteria is proposed.

Keywords: multi-objective optimization, distribution of the order volume, criterion of optimality, relative priorities of criteria.

REFERENCES

1. Guedas B., Gandibleux X., Depince Ph. Compromise Based Evolutionary Multiobjective Optimization Algorithm for Multidisciplinary Optimization, Lecture Notes in Economics and Mathematical Systems, 2011, Volume 648, pp. 69–78.
2. Seydel J., Multicriteria Support for Construction Bidding, *Mathematical and Computer Modelling*, 2001, No. 34, pp. 677–702.

УДК 62–50:519.7/8

Левин В. И.

Д-р техн. наук, профессор, Пензенский государственный технологический университет, Россия

УСТОЙЧИВОСТЬ РЕШЕНИЯ ЗАДАЧ ОПТИМАЛЬНОГО ПРОЕКТИРОВАНИЯ СИСТЕМ С ИНТЕРВАЛЬНЫМИ ПАРАМЕТРАМИ

Рассмотрена задача оптимизации неполностью определенных (недетерминированных) функций, т.е. функций с параметрами, заданными лишь с точностью до интервала. Показано, что решение этой проблемы требует также рассмотрения задачи определения устойчивости оптимума к варьированию значений параметров функции. Предлагается метод нахождения оптимума функций и определения его устойчивости методами интервальной математики.

Ключевые слова: оптимизация систем, неопределенность, устойчивость оптимума, варьирование параметров, интервальная математика.

ВВЕДЕНИЕ

На сегодняшний день в мире имеется обширная литература по оптимизации (оптимальному проектированию) различных систем с детерминированными параметрами – технических, экономических и т. д. Соответствующие задачи формулируются как задачи математического программирования с целевыми функциями и функциями ограничений, параметры которых являются детерминированными величинами. Однако на практике по объективным причинам чаще встречаются системы с недетерминированными параметрами. Оптимизация такого рода систем формализуется в виде задач математического программирования с целевыми функциями и функциями ограничений, параметры которых – различные недетерминированные величины: случайные, нечеткие, интервальные и т. д. Эти задачи, вообще говоря, сложнее детерминированных. Они требуют обобщения понятия экстремума функции, выяснения условия его существования, связанных с недетерминированностью параметров функции, и создания специальных методов поиска экстремума таких функций.

Известно три различных подхода к решению недетерминированных задач математического программирования: детерминированный, вероятностный [1] и интервальный [2]. Детерминированный подход заключается в решении задачи для определенных значений ее параметров, выбранных внутри соответствующих заданных областей неопределенности. Например, могут быть выбраны центры (середины) областей неопределенности параметров (центральная стратегия), наихудшее сочетание значений параметров задачи (пессимистическая стратегия), их наилучшее сочетание (оптимистическая стратегия) и т. д. Вероятностный подход заключается в решении задачи для усредненных (ожидаемых, в смысле математического ожидания) значений ее параметров, что предполагает задание вероятностной меры внутри соответствующих областей неопределенности. Оба указанных подхода объединяет предварительная детерминизация параметров задачи, выполняемая перед ее оптимальным решением. В отличие от них, интервальный подход не предполагает детерминизации параметров задачи, которые задаются в интервальной форме – в данном подходе оптимальное решение задачи проводится в ее «естественной форме», т.е. на основе прямого сравнения недетерминированных значений целевой функции, соответствующих различным значениям вектора аргументов, и выборе оптимального (максимального или минимального) значения данной функции. Достоинства и недостатки указанных трех подходов рассмотрены в [1–8].

Изложенные подходы к решению недетерминированных задач математического программирования, при всем их очевидном различии, объединяет одна существенная черта. А именно, все они предназначены для решения задач оптимизации, в которых параметры целевых функций и функций ограничений точно не известны. Поэтому мы не можем ограничиться простым отысканием оптимального решения нашей задачи, используя один из упомянутых выше методов. В самом деле, из-за отсутствия при решении задачи точных значений ее параметров может оказаться, что действительные значения параметров задачи несколько отличаются от тех, которые были приняты в процессе отыскания решения. В этом случае, для того чтобы найденное оптимальное решение задачи имело содержательный прикладной смысл, нам нужно, чтобы оно еще обладало следующим свойством: при небольшом варьировании значении параметров решаемой задачи ее оптимальное решение должно по-прежнему существовать. При этом точка, в которой достигается оптимум целевой функции, может переместиться из исходного положения в новое положение, которое, однако, должно быть близко к исходному. Другими словами, требуется, чтобы найденное оптимальное решение неполностью определенной (недетерминированной) задачи математического программирования было устойчивым относительно небольших количественных изменений ее параметров.

1 ПОСТАНОВКА ЗАДАЧИ

Рассмотрим сначала детерминированный случай. Пусть задана некоторая произвольная непрерывная функция n переменных

$$y = F(x_1,...,x_n), \quad (1)$$

где параметры (коэффициенты) ее явного представления $p_k, k=\overline{1,l}$, известны точно. Будем рассматривать функцию (1) в ограниченной области, определяемой системой ограничений

$$\Phi_i(x_1,...,x_n) \le b_i, \quad i=\overline{1,m}\}, \quad (2)$$

в которой параметры $q_s, s=\overline{1,t}$, явного представления функций ограничений Φ_i и правые части b_i также известны точно.

Тогда относительно функции (1) можно сформулировать полностью определенную задачу условной оптимизации (задачу математического программирования)

$$F(x_1,...,x_n) = \max, \quad (3)$$

при условии

$$\Phi_i(x_1,...,x_n) \le b_i, i=\overline{1,m}\}. \quad (4)$$

Решением задачи (3), (4) является некоторая точка $x^* = (x_1^*,...,x_n^*)$ (множество точек $M = \{x^*\}$) области (4), в которой целевая функция F достигает максимального значения $F_{\max}$. В современном математическом программировании разработано множество различных методов эффективного решения задач вида (3), (4), ориентирующихся на тип целевой функции F, а также функций ограничений $\Phi_i, i=\overline{1,m}$.

Предположим теперь, что в задаче оптимизации (3), (4) параметры явного представления целевой функции F, а также параметры явного представления функций ограничений Φ_i и правые части ограничений b_i известны не точно, а приближенно. Тогда, в соответствии со сказанным в п. 1, мы должны совместно с задачей условной оптимизации (3), (4) рассматривать еще следующую задачу: проверка устойчивости (неустойчивости) решения задачи (3), (4) относительно небольших количественных изменений ее параметров.

В отличие от существующих сегодня методов изучения устойчивости решения задач оптимизации [5], будем рассматривать все возможные количественные изменения каждого параметра задачи как единое целое. Такое рассмотрение позволяет задавать все возможные количественные изменения параметров задач оптимизации в теоретико-множественных терминах. Простейший способ такого задания заключается в том, чтобы задать совокупность указанных изменений параметров задачи в виде соответствующих числовых интервалов. Преимущество такого подхода к изучению устойчивости решения задач оптимизации состоит в том, что в его рамках изучать устойчивость задач оптимизации можно с помощью хорошо разработанных методов интервальной математики [9].

Итак, совместно с полностью определенной задачей (3), (4) мы должны рассмотреть производную от нее интервальную задачу условной оптимизации

$$\tilde{F}(x_1,...,x_n) = \max, \quad (5)$$

при условии

$$\tilde{\Phi}_i(x_1,...,x_n) \le \tilde{b}_i, i=\overline{1,m}\}. \quad (6)$$

Целевая функция $\tilde{F}$ интервальной задачи оптимизации (5), (6) получается из целевой функции F искомой, полностью определенной задачи оптимизации (3), (4) путем замены ее точных параметров $p_k, k=\overline{1,l}$, соответствующими интервальными параметрами $\tilde{p}_k = [p_{k1}, p_{k2}], k=\overline{1,l}$, которые и определяют интервальную целевую функцию $\tilde{F}$. Аналогично этому, любая функция ограничений $\tilde{\Phi}_i, i=\overline{1,m}$, интервальной задачи условной оптимизации (5), (6) получается из соответствующей функции $\Phi_i, i=\overline{1,m}$, исходной полностью определенной задачи (3), (4) заменой ее точно известных параметров $q_{si}, s=\overline{1,t}, i=\overline{1,m}$, соответствующими интервальными параметрами $\tilde{q}_{si} = [q_{si1}, q_{si2}], s=\overline{1,t}, i=\overline{1,m}$. Точно так же интервальные параметры $\tilde{b}_i, i=\overline{1,m}$, в ограничениях интервальной задачи условной оптимизации (5), (6) заменяют собой соответствующие точно известные параметры $b_i, i=\overline{1,m}$ в ограничениях исходной, полностью определенной задачи оптимизации (3), (4).

Будем называть полностью определенную задачу условной оптимизации (математического программирования) (3), (4) макроустойчивой, если она имеет решение и, кроме того, имеет решение производная от нее интервальная задача оптимизации (5), (6).

Далее, будем называть полностью определенную задачу условной оптимизации (математического программирования) (3), (4) микроустойчивой, если она макроустойчива и, сверх того, существует пара решений (x', x''), где $x' = (x_1',...,x_n')$ – некоторая точка решения задачи оптимизации (3), (4), а $x'' = (x_1'',...,x_n'')$ – некоторая точка решения задачи (5), (6), расстояние между которыми $D(x', x'')$ не превосходит заданной достаточно малой величины d.

Задача настоящего исследования – разработать алгоритмы определения макро- и микроустойчивости полностью определенных задач условной оптимизации типа (3), (4).

2 МАТЕМАТИЧЕСКИЙ АППАРАТ

В основу решения поставленной задачи положим аппарат интервальной математики [9], где алгебраические операции над интервальными числами

$\tilde{a}=[a_1,a_2]$, $\tilde{b}=[b_1,b_2]$,... вводятся как следующие теоретико-множественные конструкции

$$\tilde{a}+\tilde{b}=\{a+b \mid a\in\tilde{a}, b\in\tilde{b}\},\ \tilde{a}-\tilde{b}=$$

$$=\{a-b \mid a\in\tilde{a}, b\in\tilde{b}\},\ k\tilde{a}=\{ka \mid a\in\tilde{a}\},... \qquad (7)$$

и т. д. Другими словами, любая операция над интервалами определяется на основе соответствующей операции над точечными величинами, при условии, что конкретные значения величин пробегают все возможные значения из соответствующих интервалов. Из введенных алгебраических операций над интервалами вытекают простые правила выполнения операций:

$$[a_1,a_2]+[b_1,b_2]=[a_1+b_1,\ a_2+b_2],$$

$$[a_1,a_2]-[b_1,b_2]=[a_1-b_2,\ a_2-b_1];$$

$$k[a_1,a_2]=\begin{cases}[ka_1,ka_2],\ k>0,\\ [ka_2,ka_1],\ k<0;\end{cases}$$

$$[a_1,a_2]\cdot[b_1,b_2]=[\min_{i,j}(a_i\cdot b_j),\max_{i,j}(a_i\cdot b_j)];$$

$$[a_1,a_2]/[b_1,b_2]=[a_1.a_2]\cdot[1/b_2,1/b_1]. \qquad (8)$$

Введем теперь операции сравнения интервальных чисел [2, 8]. Попытаемся сравнить интервалы $\tilde{a}=[a_1,a_2]$ и $\tilde{b}=[b_1,b_2]$, рассматривая их как интервальные числа. Естественно начать со сравнения интервалов $\tilde{a}$ и $\tilde{b}$ на базе сравнений в отдельных парах вещественных чисел (a_i,b_j), где $a_i\in\tilde{a}, b_i\in\tilde{b}$. Однако такой подход приведет нас к провалу, так как в общем случае одни пары чисел (a_i,b_j) будут находиться в отношении $a_i>b_j$, а другие – в противоположном отношении: $a_i<b_j$. Единственное, что остается – реализовать операцию сравнения интервалов на теоретико-множественном уровне, подобно алгебраическим операциям над интервалами (7). В соответствии со вышесказанным, введем операции взятия максимума $\vee$ и минимума $\wedge$ двух интервальных чисел $\tilde{a}=[a_1,a_2]$ и $\tilde{b}=[b_1,b_2]$ в виде конструкций

$$\tilde{a}\vee\tilde{b}=\{a\vee b \mid a\in\tilde{a}, b\in\tilde{b}\},$$

$$\tilde{a}\wedge\tilde{b}=\{a\wedge b \mid a\in\tilde{a}, b\in\tilde{b}\}. \qquad (9)$$

Операция взятия максимума (минимума) из двух интервалов $\tilde{a}$ и $\tilde{b}$, согласно (9), определяется как нахождение максимума (минимума) из двух точечных величин a и b, при условии, что конкретные значения этих величин пробегают все возможные значения соответственно из интервалов $\tilde{a}$ и $\tilde{b}$. Чтобы интервалы $\tilde{a}$ и $\tilde{b}$ можно было сравнить по величине, установив их отношение ($\tilde{a}\geq\tilde{b}$ или $\tilde{a}\leq\tilde{b}$), нужно, чтобы: 1) введенные операции $\vee,\wedge$ над этими интервалами существовали; 2) их результатом был один из операндов $\tilde{a}$ или $\tilde{b}$; 3) операции $\vee,\wedge$ были согласованными между собой, т. е. было выполнено условие: если большим (меньшим) является один из интервалов $\tilde{a},\tilde{b}$, то меньшим (большим) является другой из них. Условие сравнимости величин двух интервалов является, очевидно, необходимым и достаточным условием. Однако легко доказать, что условие согласованности операций $\vee$ и $\wedge$ над интервалами выполняется всегда (для любой пары интервалов $(\tilde{a},\tilde{b})$). Также всегда (для любой пары интервалов) выполняется условие существования введенных нами выше операций взятия максимума $\vee$ и минимума $\wedge$ двух интервалов, причем результатом операции оказывается некоторый, вообще говоря, новый интервал. В конечном итоге необходимое и достаточное условие сравнимости интервалов $\tilde{a}$ и $\tilde{b}$ превращается в условие, по которому операции $\tilde{a}\vee\tilde{b}$ и $\tilde{a}\wedge\tilde{b}$ должны давать в результате обязательно один из интервалов-операндов: $\tilde{a}$ или $\tilde{b}$. Такая формулировка условия сравнимости интервалов дает возможность получения его в конструктивной форме, пригодной, к тому же, для практического применения. Это является базовой формой условия.

Теорема 1. Для сравнимости двух интервалов $\tilde{a}=[a_1,a_2]$ и $\tilde{b}=[b_1,b_2]$ и их нахождения между собой в отношении $\tilde{a}\geq\tilde{b}$ необходимо и достаточно, чтобы одноименные границы этих интервалов удовлетворяли условиям

$$a_1\geq b_1,\ a_2\geq b_2, \qquad (10)$$

а для сравнимости этих интервалов и их нахождения между собой в отношении $\tilde{a}\leq\tilde{b}$ – чтобы удовлетворялись следующие условия:

$$a_1\leq b_1,\ a_2\leq b_2. \qquad (11)$$

Согласно утверждению теоремы 1, интервалы $\tilde{a}$ и $\tilde{b}$ являются сравнимыми и находятся в определенном отношении порядка $\tilde{a}\geq\tilde{b}$ или $\tilde{a}\leq\tilde{b}$ только когда в таком же отношении находятся их одноименные границы a_1,b_1 и a_2,b_2. Другими словами, для сравнимости интервалов меньший интервал должен быть сдвинут обеими границами влево относительно большего интервала. Итак, с помощью теоремы 1 сравнение двух интервалов и выбор большего (меньшего) из них сводится к сравнению одноименных границ этих интервалов, являющихся точно известными вещественными числами.

Теорема 2. Для несравнимости двух интервалов $\tilde{a}=[a_1,a_2]$ и $\tilde{b}=[b_1,b_2]$, т. е. для того, чтобы они не находились ни в отношении $\tilde{a}\geq\tilde{b}$, ни в отношении $\tilde{a}\leq\tilde{b}$, необходимо и достаточно, чтобы одноименные границы интервалов удовлетворяли условиям

$$a_1<b_1,\ a_2>b_2 \quad \text{или} \quad b_1<a_1,\ b_2>a_2. \qquad (12)$$

Стоит отметить, что условия (12) обозначают ту ситуацию, когда один интервал на числовой оси полностью «накрывает» другой.

Теорема 2 показывает существование случаев несравнимости интервалов. Несравнимость некоторых интервалов – естественное следствие того, что, в отличие от точных вещественных чисел, интервальные числа задаются с некоторой неопределенностью (точно известно, что вещественное число принимает некоторое значение в заданном интервале, но не известно, какое именно это значение). Далее, теоремы 1 и 2, посвященные сравнению пар интервалов, можно обобщить на системы с произвольным числом интервалов.

Теорема 3. Для существования в системе интервалов $\tilde{a}(1)=[a_1(1),a_2(1)], \tilde{a}(2)=[a_1(2),a_2(2)],...$ максимального интервала необходимо и достаточно, чтобы его границы располагались относительно одноименных границ всех остальных интервалов согласно следующим условиям

$$a_1(1) \ge a_1(2),\ a_1(1) \ge a_1(3),...;$$

$$a_2(1) \ge a_2(2),\ a_2(1) \ge a_2(3),... \qquad (13)$$

Условия-неравенства (13) записаны для конкретного случая, когда максимальным является интервал $\tilde{a}(1)$, что не ограничивает общности.

Теорема 4. Для существования в системе интервалов $\tilde{a}(1)=[a_1(1),a_2(1)], \tilde{a}(2)=[a_1(2),a_2(2)],...$ минимального интервала необходимо и достаточно, чтобы его границы были расположены относительно одноименных границ всех остальных интервалов согласно условиям

$$a_1(1) \le a_1(2),\ a_1(1) \le a_1(3),...;$$

$$a_2(1) \le a_2(2),\ a_2(1) \le a_2(3),... \qquad (14)$$

Условия (14), аналогично условиям (13), записаны для случая, когда минимальным является интервал $\tilde{a}(1)$, что не ограничивает общности.

Теоремы 3, 4 означают, что интервал является максимальным (минимальным) из интервалов системы только если максимальны (минимальны) его нижняя граница – среди нижних границ всех интервалов – и верхняя граница – среди верхних границ всех интервалов.

3 МАКРОУСТОЙЧИВОСТЬ ЗАДАЧИ УСЛОВНОЙ ОПТИМИЗАЦИИ

Обратимся к полностью определенной задаче условной оптимизации (3), (4) и опишем метод установления макроустойчивости этой задачи. Полностью определенная задача условной оптимизации (3), (4) по определению (см. п. 2) является макроустойчивой, если она сама и производная от нее интервальная задача условной оптимизации (5), (6) имеют решения. Существование решения полностью определенной задачи условной оптимизации (3), (4) можно установить с помощью общеизвестных методов математического программирования [10–12], так что здесь нет никаких проблем. Сложнее обстоит дело с проверкой существования решения интервальной задачи условной оптимизации (5), (6). Здесь эффективным оказывается применение детерминизационного метода решения задач интервальной оптимизации [2, 8, 13].

Интервальная задача условной оптимизации (5), (6) имеет интервальную целевую функцию $\tilde{F}(x_1,...,x_n)$, интервальные функции ограничений $\tilde{\Phi}_i, \overline{1,m}$, в левых частях ограничений и интервальные параметры $\tilde{b}_i, i=\overline{1,m}$, в правых частях. Используя формулы элементарных преобразований интервалов (8), функции $\tilde{F}$ и $\tilde{\Phi}_i$ можно представить явно в интервальной форме. Так же можно представить и параметры $\tilde{b}_i$. Все эти представления записываются в виде

$$\begin{aligned}
&\tilde{F}(x_1,...,x_n)=[F_1(x_1,...,x_n),F_2(x_1,...,x_n)],\\
&\tilde{\Phi}_i(x_1,...,x_n)=[\Phi_{i1}(x_1,...,x_n),\Phi_{i2}(x_1,...,x_n)],\ i=\overline{1,m},\\
&\tilde{b}_i=[b_{i1},b_{i2}],\qquad\qquad\qquad\qquad i=\overline{1,m}.
\end{aligned} \qquad (15)$$

После этого всю интервальную задачу условной оптимизации (5), (6) также можно переписать в явном интервальном виде

$$[F_1(x_1,...,x_n),F_2(x_1,...,x_n)] = \max, \qquad (16)$$

$$[\Phi_{i1}(x_1,...,x_n),\Phi_{i2}(x_1,...,x_n)] \le [b_{i1},b_{i2}],\ \ i=\overline{1,m}\ \}. \qquad (17)$$

От интервального представления задачи (16), (17) перейдем к ее эквивалентному представлению в виде пары полностью определенных (детерминированных) задач условной оптимизации, которое уже поддается решению. Для этого сначала по теореме 3 представим интервальное уравнение (16) в виде эквивалентной пары детерминированных уравнений

$$F_1(x_1,...,x_n)=\max, F_2(x_1,...,x_n)=\max. \qquad (18)$$

Далее, по теореме 1 представим систему интервальных неравенств (17) в виде эквивалентной системы обычных детерминированных неравенств

$$\Phi_{i1}(x_1,...,x_n) \le b_{i1},\ \Phi_{i2}(x_1,...,x_n) \le b_{i2},\ i=\overline{1,m}. \qquad (19)$$

Соединив пару уравнений оптимизации (18) с системой неравенств-ограничений (19), получаем совокупность двух полностью определенных задач условной оптимизации вида (3), (4)

$$\left.\begin{aligned}
&F_1(x_1,...,x_n)=\max,\\
&\Phi_{i1}(x_1,...,x_n) \le b_{i1},\ i=\overline{1,m},\\
&\Phi_{i2}(x_1,...,x_n) \le b_{i2},\ i=\overline{1,m},
\end{aligned}\right\} \qquad (20)$$

$$\left.\begin{aligned} &F_2(x_1,...,x_n) = \max, \\ &\Phi_{i1}(x_1,...,x_n) \le b_{i1},\ i=\overline{1,m}, \\ &\Phi_{i2}(x_1,...,x_n) \le b_{i2},\ i=\overline{1,m}, \end{aligned}\right\} \qquad (21)$$

эквивалентную исходной интервальной задаче условной оптимизации (5), (6). Задачу (20) будем называть нижней граничной задачей интервальной задачи (5), (6), а задачу (21) – верхней граничной задачей. Итак, для получения решения интервальной задачи (5), (6) нужно решить ее нижнюю (20) и верхнюю (21) граничные задачи. В общем случае решение нижней граничной задачи имеет вид $\{M_{\text{н}}(x), F_{1,\max}\}$, а верхней граничной задачи – вид $\{M_{\text{в}}(x), F_{2,\max}\}$.

Здесь $M_{\text{н}}(x)$, $M_{\text{в}}(x)$ – множества точек решения $x=(x_1,...,x_n)$ нижней и верхней граничных задач, а $F_{1,\max}$, $F_{2,\max}$ – полученные максимальные значения целевых функций этих задач. Решение интервальной задачи оптимизации (5), (6) формируется из решений ее нижней и верхней граничных задач и имеет вид

$$\{x^* \in M_{\text{н}}(x) \cap M_{\text{в}}(x);\quad \tilde{F}_{\max} = [F_{1,\max}, F_{2,\max}]\ \}. \qquad (22)$$

Согласно (22), в качестве точки решения x^* интервальной задачи оптимизации (5), (6) выбирается любая точка из пересечения множеств точек решения ее нижней и верхней граничных задач, а в качестве максимального значения интервальной целевой функции $\tilde{F}_{\max}$ – интервал от максимального значения целевой функции нижней задачи $F_{1,\max}$ до максимального значения целевой функции верхней задачи $F_{2,\max}$.

Из выполненного процесса построения решения интервальной задачи условной оптимизации вида (5), (6) и определения макроустойчивости полностью определенной задачи условной оптимизации (3), (4) вытекает следующая основная теорема.

Теорема 5. Для того чтобы полностью определенная задача условной оптимизации (3), (4) была макроустойчива, необходимо и достаточно, чтобы: 1) эта задача имела решение; 2) интервальная задача оптимизации (5), (6), производная от задачи (3), (4), имела нижнюю и верхнюю граничные задачи, обладающие решениями; 3) множества решений нижней и верхней граничных задач интервальной задачи оптимизации (5), (6) пересекались.

Сформулированная выше теорема 5 определяет следующий алгоритм для проверки произвольной полностью определенной (детерминированной) задачи условной оптимизации, имеющей вид (3), (4), на макроустойчивость:

Шаг 1. Используя подходящие для конкретного типа целевой функции методы решения полностью определенных (детерминированных) задач условной оптимизации [10–12], ищем решение $x'=(x'_1,...,x'_n)$ задачи (3), (4). Одновременно с этим проверяется и существование (несуществование) решения этой задачи.

Шаг 2. Задаваясь некоторыми подходящими значениями интервальных параметров целевой функции F, функций ограничений Φ_i, $i=\overline{1,m}$, и правых частей ограничений b_i, $i=\overline{1,m}$ полностью определенной задачи условной оптимизации (3), (4), строим производную от нее интервальную задачу условной оптимизации (5), (6).

Шаг 3. Используя формулы интервальной математики (8), выражающие результаты элементарных преобразований интервалов, представляем целевую функцию $\tilde{F}$, функции ограничений $\tilde{\Phi}_i$, $i=\overline{1,m}$, а также правые части ограничений $\tilde{b}_i$, $i=\overline{1,m}$, интервальной задачи условной оптимизации (5), (6) в интервальной форме (15).

Шаг 4. По найденным на шаге 3 интервальным представлениям функций $\tilde{F}, \tilde{\Phi}_i$, $i=\overline{1,m}$, и параметров $\tilde{b}_i$, $i=\overline{1,m}$, формируем нижнюю (20) и верхнюю (21) граничные задачи интервальной задачи оптимизации (5), (6).

Шаг 5. Используя те же самые методы, что и на шаге 1, ищем решения оптимизационных задач (20) и (21). Одновременно с этим проверяем существование или несуществование решений указаных задач. Полные решения задач условной оптимизации вида (20, (21) имеют соответственно вид $\{M_{\text{н}}(x), F_{1,\max}\}, \{M_{\text{в}}(x), F_{2,\max}\}$, где $M_{\text{н}}(x)$ – множество точек x решения нижней, $M_{\text{в}}(x)$ – множество точек x решения верхней граничной задачи.

Шаг 6. Проверяется наличие (отсутствие) пересечения найденных в результате решения задач (20) и (21) множеств $M_{\text{н}}(x), M_{\text{в}}(x)$.

Итог. Если в результате работы алгоритма выяснилось, что полностью определенная задача условной оптимизации (3), (4) имеет решение, а производная от нее интервальная задача (5), (6) имеет нижнюю и верхнюю граничные задачи, обладающие решениями, причем множества этих решений пересекаются, то задача оптимизации (3), (4) является макроустойчивой. В противном случае задача (3), (4) не является макроустойчивой.

4 МИКРОУСТОЙЧИВОСТЬ ЗАДАЧИ УСЛОВНОЙ ОПТИМИЗАЦИИ

Снова обратимся к полностью определенной задаче условной оптимизации (3), (4) и опишем метод установления микроустойчивости этой задачи.

Полностью определенная задача условной оптимизации (3), (4) по определению (см. п. 2) является микроустойчивой, если она обладает свойством макроустойчивости и, кроме того, существует пара решений (x', x''), где $x'=(x'_1,...,x'_n)$ – некоторое решение исходной полностью определенной задачи (3), (4), а $x''=(x''_1,...,x''_n)$ – какое-то решение производной от нее интервальной задачи условной оптимизации (5), (6), расстояние между которыми $D(x', x'')$ не превосходит заданной достаточно малой величины d. Из этого определения напрямую вытекает следующий алгоритм проверки произвольной полностью определенной задачи условной оптимизации вида (3), (4) на микроустойчивость.

Шаг 1. С помощью 6-шагового алгоритма, изложенного в п. 4, проверяем задачу (3), (4) на макроустойчивость. В случае отрицательного результата (задача (3), (4) не макроустойчива) конец алгоритма, с выводом: задача (3), (4) не является микроустойчивой. При положительном результате проверки (задача (3), (4) макроустойчива) переход к шагу 2.

Шаг 2. Выбираем некоторую произвольную точку решения $x' = (x'_1, ..., x'_n)$ задачи (3), (4), найденную на шаге 1. После этого добавляем к ней какую-либо точку решения $x'' = (x''_1, ..., x''_n)$ соответствующей интервальной задачи (5), (6), также найденную на шаге 1. В результате получаем пару решений (x', x'') указанных двух задач.

Шаг 3. Вычисляем величину расстояния $D(x', x'')$ между точками решения x', x'' указанных двух задач, используя для этого формулу

$$D(x', x'') = \sqrt{(x'_1 - x''_1)^2 + ... + (x'_n - x''_n)^2} \ . \qquad (23)$$

Шаг 4. Проверяем выполнение неравенства, сравнивающего расстояние $D(x', x'')$ с некоторой изначально заданной достаточно малой величиной d:

$$D(x', x'') \le d \ . \qquad (24)$$

Если условие (24) выполнено, задача оптимизации (3), (4) объявляется микроустойчивой и конец алгоритма. В противном случае совершается переход к шагу 2, в котором теперь к точке решения $x' = (x'_1, ..., x'_n)$ задачи (3), (4), найденной на шаге 1, добавляется какая-то другая точка решения $x'' = (x''_1, ..., x''_n)$ задачи (5), (6) из числа найденных на шаге 1. В результате получаем новую пару решений (x', x'') и т. д.

Итог. Если в результате работы алгоритма после некоторого достаточного числа шагов получена пара решений (x', x''), удовлетворяющая неравенству (24), процедура останавливается и задача (3), (4) объявляется микроустойчивой. В противном случае процедура также останавливается, но задача (3), (4) признается не обладающей свойством микроустойчивости.

ЗАКЛЮЧЕНИЕ

В статье показано, что проблема оптимизации неполностью определенных функций не может ограничиться только отысканием точки оптимума и значения в ней нашей функции, но и должна включать в себя задачу определения устойчивости найденного оптимума. Последнее означает, что при небольшом варьировании параметров оптимизируемой функции ее оптимум должен по-прежнему существовать и находиться в точке, близкой к точке исходного оптимума. Для установления устойчивости оптимума неполностью определенных функций предложена специальная эффективная методика, основанная на аппарате интервальной математики.

СПИСОК ЛИТЕРАТУРЫ

1. *Первозванский, А. А.* Математические модели в управлении производством / А. А. Первозванский. – М. : Наука, 1975. – 616 с.
2. *Левин, В. И.* Интервальное дискретное программирование / В. И. Левин // Кибернетика и системный анализ. – 1994. – № 6. – С. 91–103.
3. *Libura, M.* Integer Programming Problems with Inexact Objective Function / M. Libura // Control and Cybernetic. – 1980. – Vol. 9, No. 4. – P. 189–202.
4. *Тимохин, С. Г.* О задачах линейного программирования в условиях неточных данных / С. Г. Тимохин, А. В. Шапкин // Экономика и математические методы. – 1981. – Том 17, № 5. – С. 955–963.
5. *Рощин, В. А.* Вопросы решения и исследования одного класса задач неточного целочисленного программирования / В. А. Рощин, Н. В. Семенова, И. В. Сергиенко // Кибернетика. – 1989. – № 2. – С. 42–46.
6. *Семенова, Н. В.* Решение одной задачи обобщенного целочисленного программирования / Н. В. Семенова // Кибернетика. – 1984. – № 5. – С. 25–31.
7. *Вощинин, А. П.* Оптимизация в условиях неопределенности / А. П. Вощинин, Г. Р. Сотиров. – М. : Изд-во МЭИ, 1989. – 224 с.
8. *Левин, В. И.* Интервальные методы оптимизации систем в условиях неопределенности / *В. И. Левин.* – Пенза : изд-во Пензенского технологического института, 1999. – 95 с.
9. *Алефельд, Г.* Введение в интервальные вычисления / Г. Алефельд, Ю. Херцбергер. – М. : Мир, 1987. – 360 с.
10. *Юдин, Д. Б.* Задачи и методы линейного программирования / Д. Б. Юдин, Е. Г. Гольдштейн. – М. : Советское радио, 1964. – 350 с.
11. *Корбут, А. А.* Дискретное программирование / А. А. Корбут, Ю. Ю. Финкельштейн. – М. : Наука, 1969. – 280 с.
12. *Левин, В. И.* Структурно-логические методы исследования сложных систем / В. И. Левин. – М. : Наука, 1987. – 304 с.
13. *Левин, В. И.* Дискретная оптимизация в условиях интервальной неопределенности / В. И. Левин // Автоматика и телемеханика. – 1992. – № 7. – С. 97–106.

Стаття надійшла до редакції 13.08.2013.

Левін В. І.
Д-р техн. наук, професор, Пензенській державний технологічний університет, Росія

СТАБІЛЬНІСТЬ ВИРІШЕННЯ ЗАВДАНЬ ОПТИМАЛЬНОГО ПРОЕКТУВАННЯ СИСТЕМ З ІНТЕРВАЛЬНИМИ ПАРАМЕТРАМИ

Розглянуто задачу оптимізації неповністю певних (недетермінованих) функцій, тобто функцій з параметрами, заданими лише з точністю до інтервалу. Показано, що вирішення цієї проблеми потребує також розгляду завдання визначення стійкості оптимуму до варіювання значень параметрів функції. Пропонується метод знаходження оптимуму функцій і визначення його стійкості методами інтервальної математики.

Ключові слова: оптимізація систем, невизначеність, стійкість оптимуму, варіювання параметрів, інтервальна математика.

Levin V. I.
Dr. Sci., professor, Penza State Technological University, Russia

THE STABILITY OF SOLUTION OF SYSTEMS WITH UNDEFINED PARAMETERS OPTIMAL DESIGN PROBLEM

Our article considers the problem of optimization of incompletely specified functions, namely, functions which parameters are given within range of possible values. It is shown that solution of this problem also requires solving problem of determining stability of optimum of such functions to variation of values of their parameters. A method for obtaining such optimum of incompletely defined functions is presented. Method uses determination of problem. It allows to split original non-deterministic problem into two optimization problem of deterministic functions, which are solved separately. After that solutions are combined into one which is a solution of original problem. The article also provides a method for determining the stability of the optimum found of incompletely defined functions by methods of interval mathematics. We formulate 5 theorems determining the conditions for existence of optimum of incompletely defined function and its resistance to changing the function parameters. Algorithms for verifying the stability function are given

Keywords: system optimization, uncertainty, stability of optimum, variation of parameters, interval mathematics.

REFERENCES

1. Pervozvanskiy A. A. Matematicheskie modeli v upravlenii proizvodstvom, Moscow, Nauka, 1975, 616 p.
2. Levin V. I. Intervaljnoe diskretnoe programmirovanie, *Kibernetika i sistemniy analiz*, 1994, No. 6, pp. 91–103.
3. Libura M. Integer Programming Problems with Inexact Objective Function, *Control and Cybernetic*, 1980, Vol. 9, No. 4, pp. 189–202.
4. Timokhin S. G., Shapkin A. V. O zadachah linejnogo programmirovaniya v usloviyah netochnyh dannyh, *Ekonomika i matematicheskie metody*, 1981, Vol. 17, No. 5, pp. 955–963.
5. Rothin V. A., Semenova N. V., Sergienko I. V. Voprosy resheniya i issledovaniya odnogo klassa zadach netochnogo celochislennogo programmirovaniya, *Kibernetika*, 1989, No. 2, pp. 42–46.
6. Semenova N. V. Reshenie odnoy zadachi obobschennogo celochislennogo programmirovaniya, *Kibernetika*, 1984, No. 5, pp. 25–31.
7. Voschinin A. P., Sotirov G. R. Optimizaciya v usloviyah neopredelennosti, Moscow, Izd-vo MEI, 1989, 224 p.
8. Levin V. I. Intervalnie metody optimizacii sistem v usloviyah neopredelennosti. Penza, Izd-vo Penzenskogo tehnologicheskogo instituta, 1999, 95 p.
9. Alefeld G., Hercberger Yu. Vvedenie v intervalnye vychisleniya. Moscow, Mir, 1987, 360 p.
10. Yudin D. B., Goldshtein E. G. Zadachi i metody linejnogo programmirovaniya. Moscow, Sovetskoe radio, 1964, 350 p.
11. Korbut A. A., Finkelshtein Yu. Yu. Diskretnoe programmirovanie. Moscow, Nauka, 1969, 280 p.
12. Levin V. I. Strukturno-logicheskie metody issledovaniya slozhnyh sistem. Moscow, Nauka, 1987, 304 p.
13. Levin V. I. Diskretnaya optimizaciya v usloviyah intervalnoy neopredelennosti, *Avtomatika i telemekhanika*, 1992, No. 7, pp. 97–106.

УДК 519.1

Нікуліщев Г. І.

Старший викладач, Запорізький національний технічний університет, Україна, E-mail: tymahod@ukr.net

ПРОТОКОЛ СЛІПОГО ЕЛЕКТРОННОГО ЦИФРОВОГО ПІДПИСУ НА ЕЛІПТИЧНИХ КРИВИХ НАД СКІНЧЕНИМ ВЕКТОРНИМ ПОЛЕМ

У даній статті пропонується реалізація протоколу сліпого електронного цифрового підпису, що являє собою модифікацію стандарту ГОСТ Р 34.10-2001 на еліптичних кривих над скінченим векторним полем. Аналізується захищеність запропонованого протоколу за критерієм анонімності.

Ключові слова: сліпий електронний цифровий підпис, скінчене векторне поле, еліптична крива.

ВСТУП

Останнім часом особливої актуальності набуло питання забезпечення чесності та прозорості процедури виборів. У зв'язку з цим пильної уваги наукової спільноти набули механізми електронного голосування. Одним з методів реалізації такого механізму є використання сліпого електронного цифрового підпису (ЕЦП). Оскільки на разі схеми сліпого ЕЦП не стандартизовані, широко використовуються як власноручно розроблені схеми, так і модифікації існуючих стандартів.

В статті розглядається модифікація російського стандарту ЕЦП ГОСТ Р 34.10-2001 «Информационная технология. Криптографическая защита информации. Процессы формирования и проверки ЭЦП». Цей стандарт базується на математичному апараті еліптичних кривих (ЕК) над простим полем Галуа. Для забезпечення необхідної криптостійкості стандартом рекомендується використовувати параметри алгоритму розміром 256 біт і вище, що зобумовлює достатню складність групової операції в групі точок ЕК. Як варіант оптимізації групової операції можна розглядати використання ЕК над скінченим векторним полем (СВП), яке дозволяє зберегти необхідний порядок групи точок ЕК при меншому розмірі елементів поля [1].

Сліпий ЕЦП вирішує специфічну задачу підтвердження справжності документів без розкриття їхнього авторства і, завдяки цьому, може використовуватись в схемах електронного голосування. В алгоритмі сліпого ЕЦП один учасник формує документ, а інший підписує його всліпу без можливості ознайомитися із вмістом. При цьому важливо, щоб навіть підписувач не зміг встановити автора документа. Через це до інших критеріїв захищеності схем ЕЦП у випадку сліпого підпису додається критерій анонімності. Він показує неможливість визначити автора документа з боку підписувача, якщо він використовуватиме всі відомі йому параметри, які використовувались при постановці підпису.

1 ЕЛІПТИЧНІ КРИВІ НАД СКІНЧЕНИМ ВЕКТОРНИМ ПОЛЕМ

Розглянемо одне з можливих розширень поля Галуа $GF(p)$ – скінчене векторне поле. В СВП входять вектори певної довжини n

$$v = v_1 \cdot e_1 + v_2 \cdot e_2 + ... + v_n \cdot e_n = (v_1, v_2, ..., v_n), \qquad (1)$$

де $v_i \in GF(p)$, e_i – базисні вектори, $i = 1, 2, ..., n$.

На множині векторів визначимо операції додавання та множення.

Додавання двох векторів $u = (u_1, u_2, ..., u_n)$ та $v = (v_1, v_2, ..., v_n)$ відбувається за формулою

$$w = u + v = (u_1 + v_1 \bmod p, u_2 + v_2 \bmod p, ..., u_n + v_n \bmod p). \qquad (2)$$

Одиничним елементом за операцією додавання є вектор $u_0 = (0, 0, ..., 0)$. Відповідно, вектор $u = (u_1, u_2, ..., u_n)$ називається оберненим до вектора $v = (v_1, v_2, ..., v_n)$, якщо $u + v = u_0$, тобто

$$v = -u = (-u_1 \bmod p, -u_2 \bmod p, ..., -u_n \bmod p). \qquad (3)$$

Множення вектора v на число k реалізується за формулою

$$k \cdot a = (k \cdot a_1 \bmod p, k \cdot a_2 \bmod p, ..., k \cdot a_n \bmod p). \qquad (4)$$

Позначимо операцію множення двох векторів знаком ∘%. Ця групова операція мультиплікативної групи СВП визначається асоціативною таблицею множення базисних векторів [1], коефіцієнти перемножуються за принципом множення многочленів. В якості одиничного елемента СВП оберемо, наприклад, вектор $v_0 = (1, 0, ..., 0)$. Відповідно, два елементи СВП, результатом множення яких між собою є одиничний елемент, називаються взаємно оберненими.

Проілюструємо правило множення базисних векторів для довжини вектора $n = 2$ за допомогою табл. 1, для $n = 3$ – табл. 2. Значення розтягуючих коефіцієнтів τ і μ обчислюються з системи характеристичних рівнянь, що задає існування оберненого вектора для будь-якого заданого v.

Таблиця 1. Множення базисних векторів СВП для $n = 2$

$\circ$	e_1	e_2
e_1	e_1	e_2
e_2	e_2	$\tau \cdot e_2$

Таблиця 2. Множення базисних векторів СВП для $n = 3$

$\circ$	e_1	e_2	e_3
e_1	e_1	e_2	e_3
e_2	e_2	$\tau \cdot e_3$	$\tau \cdot \mu \cdot e_1$
e_3	e_3	$\tau \cdot \mu \cdot e_1$	$\mu \cdot e_2$

Відповідно до табл.1 система характеристичних рівнянь для $n = 2$ має вигляд

$$\begin{aligned} v_1 \cdot x + \tau \cdot v_2 \cdot y &= 1 \bmod p, \\ v_2 \cdot x + v_1 \cdot y &= 0 \bmod p. \end{aligned} \tag{5}$$

Якщо визначник $\Delta = {v_1}^2 - \tau \cdot {v_2}^2 \bmod p$ системи (5) приймає нульове значення, то для вектора $v = (v_1, v_2)$ не існує оберненого. Звідси, τ має бути квадратичним незалишком за модулем p. Тоді всі ненульові вектори СВП будуть утворювати мультиплікативну групу порядку $(p^2 - 1)$.

Відповідно до табл. 2 система характеристичних рівнянь для n=3 має вигляд

$$\begin{aligned} v_1 \cdot x + \tau \cdot \mu \cdot v_3 \cdot y + \tau \cdot \mu \cdot v_2 \cdot z &= 1 \bmod p, \\ v_2 \cdot x + v_1 \cdot y + \mu \cdot v_3 \cdot z &= 0 \bmod p, \\ v_3 \cdot x + \tau \cdot v_2 \cdot y + v_1 \cdot z &= 0 \bmod p. \end{aligned} \tag{6}$$

Якщо визначник системи (6) приймає нульове значення, то для вектора $v = (v_1, v_2, v_3)$ не існує оберненого. Величина $(p - 1)$ має ділитися на 3, а кожен з добутків $\tau^2 \cdot \mu$ та $\tau \cdot \mu^2$ має бути кубічним незалишком за модулем p. Тоді всі ненульові вектори СВП будуть утворювати мультиплікативну групу порядку $(p^3 - 1)$.

Розглянемо визначення ЕК над СВП. Оскільки характеристика СВП дорівнює p, то для $p \neq 2$ та $p \neq 3$ рівняння кривої можна записувати в скороченій формі

$$y \circ y = x \circ x \circ x + a \circ x + b, \tag{7}$$

де x, y, a, b – елементи СВП.

Кожна пара векторів x, y, що задовольняє рівнянню (7), вважається точкою ЕК. Сукупність всіх точок ЕК разом з нескінченно віддаленою точкою O складає групу точок ЕК над СВП.

Групова операція додавання точок ЕК над СВП визначається аналогічно до групової операції додавання точок ЕК над простим полем [2] з використанням операції множення векторів $\circ$. Координати точки $U = (x^U, y^U)$, яка представляє собою суму точок $S = (x^S, y^S)$ та $T = (x^T, y^T)$, визначаються за формулами

$$x^U = \lambda \circ \lambda - x^S - x^T, \tag{8}$$

$$y^U = \lambda \circ (x^S - x^U) - y^S. \tag{9}$$

Величина λ обчислюється за формулою (10), якщо $S \neq T$

$$\lambda = \frac{y^T - y^S}{x^T - x^S}, \tag{10}$$

або за формулою (11), якщо $S = T$

$$\lambda = \frac{3 \cdot x^S \circ x^S + a}{2 \cdot y^S}, \ \lambda = \frac{3 \cdot x^S \circ x^S + a}{2 \cdot y^S}, \tag{11}$$

де a – коефіцієнт ЕК.

Виконання умови гладкості

$$4 \cdot a \circ a \circ a + 27 \cdot b \circ b \neq (0, 0, ...) \tag{12}$$

забезпечує утворення групи точками ЕК.

В криптографії застосовуються гладкі ЕК з великим простим порядком групи точок. Порядок СВП дорівнює p^n, отже оцінку порядка групи точок ЕК над СВП $\#E(p)$ за теоремою Хассе [2] можна виразити наступним співвідношенням

$$p^n + 1 - 2 \cdot \sqrt{p^n} \leq \#E(p) \leq p^n + 1 + 2 \cdot \sqrt{p^n}. \tag{13}$$

Згідно з нижньою границею співвідношення (13), для досягнення порядку групи точок ЕК величиною, наприклад, 178 біт, необхідно використовувати просте число p величиною 90 біт за довжини вектора $n = 2$.

2 ПРОТОКОЛ СЛІПОГО ПІДПИСУ НА ОСНОВІ ГОСТ Р 34.10-2001

Розглянемо наступний протокол сліпого ЕЦП, запропонований в [3], який базується на стандарті ГОСТ Р 34.10-2001.

В схемі приймають участь дві сторони: підписувач А та абонент Б. Абонент Б виступає в якості емітента документа m, який підписувач А має підписати наосліп. Валідатором може виступити будь-хто з них або третя особа. Перевірка підпису відбувається за допомогою відкритого ключа підписувача А.

Загальні параметри: просте поле $GF(p)$, ЕК над цим полем з групою точок простого порядку q, базова точка P, хеш-функція $H()$.

Протокол складається з трьох етапів – генерація ключів, постановка підпису, перевірка підпису.

Під час генерації ключів секретний ключ d обирається випадково з діапазону $1 < d < q$. Відкритий ключ Q отримується з нього за формулою

$$Q = d \cdot P. \quad (14)$$

Етап постановки підпису починає підписувач А, обираючи одноразовий ключ k з діапазону $1 < k < q$ та обчислюючи точку E за формулою

$$E = k \cdot P = (x^E, y^E). \quad (15)$$

Підписувач А відправляє точку E абоненту Б.

Абонент Б формує хеш-образ повідомлення за співвідношенням

$$h = H(m). \quad (16)$$

Після цього він випадково обирає маскуючі параметри α та β з діапазону $1 < \alpha, \beta < q$ і обчислює точку C за формулою

$$C = \alpha \cdot E + \beta \cdot P = (x^C, y^C). \quad (17)$$

Абонент Б обчислює величини r та r' за формулами

$$r = x^C \bmod q, \quad (18)$$

$$r' = x^E \bmod q. \quad (19)$$

Ці величини використовуються в формулі засліплення хеш-образу повідомлення

$$h' = \frac{r'}{r} \cdot h \cdot \alpha \bmod q. \quad (20)$$

Абонент Б пересилає засліплений хеш-образ повідомлення h' підписувачу А.

Підписувач А ставить під ним засліплений підпис s' за допомогою власного секретного ключа d

$$s' = (d \cdot r' + k \cdot h') \bmod q, \quad (21)$$

та пересилає отримане значення абоненту Б.

Абонент Б має можливість перевірити справжність засліпленого підпису s' за допомогою співвідношення (22), використовуючи відкритий ключ Q підписувача А

$$s' \cdot P = r' \cdot Q + h' \cdot E. \quad (22)$$

Якщо засліплений підпис проходить перевірку, абонент Б формує з нього остаточний підпис

$$s = (s' \cdot \frac{r}{r'} + \beta \cdot h) \bmod q. \quad (23)$$

Сліпим підписом під документом m вважається пара значень $< r, s >$.

Валідатор при перевірці підпису $\{m, < r, s >\}$ обчислює точку R, використовуючи відкритий ключ Q підписувача А

$$R = \frac{s}{h} \cdot P - \frac{r}{h} \cdot Q = (x^R, y^R). \quad (24)$$

Підпис вважається справжнім, якщо виконується співвідношення

$$r = x^R \bmod q. \quad (25)$$

Автором було доведено в [4], що наведений протокол є захищеним за критерієм анонімності, тому він підходить для модифікації.

3 ПРОТОКОЛ СЛІПОГО ПІДПИСУ НА ЕК НАД СВП

Автором пропонується протокол сліпого підпису, який базується на вищенаведеному, однак використовує математичний апарат ЕК над СВП.

В схемі приймають участь дві сторони: підписувач А та абонент Б. Абонент Б виступає в якості емітента документа m, який підписувач А має підписати наосліп. Валідатором може виступити будь-хто з них або третя особа. Перевірка підпису відбувається за допомогою відкритого ключа підписувача А.

Загальні параметри: скінчене векторне поле з довжиною вектора n – розширення простого поля $GF(p)$, ЕК над цим полем з групою точок простого порядку q, базова точка P, хеш-функція $H()$.

Протокол складається з трьох етапів – генерація ключів, постановка підпису, перевірка підпису.

Під час генерації ключів секретний ключ d обирається випадково з діапазону $1 < d < q$. Відкритий ключ Q отримується з нього за формулою

$$Q = d \cdot P. \quad (26)$$

Оскільки ЕК задана над СВП, то при обчисленні суми двох точок в цій та подальших формулах використовуються співвідношення (8–11).

Етап постановки підпису починає підписувач А, обираючи одноразовий ключ k з діапазону $1 < k < q$, обчислюючи точку E за формулою

$$E = k \cdot P = (x^E, y^E) \quad (27)$$

та відправляючи цю точку абоненту Б.

Абонент Б формує хеш-образ повідомлення за співвідношенням

$$h = H(m). \quad (28)$$

Після цього він випадково обирає маскуючі параметри α та β з діапазону $1 < \alpha, \beta < q$ і обчислює точку C за формулою

$$C = \alpha \cdot E + \beta \cdot P = (x^C, y^C). \quad (29)$$

Абонент Б обчислює величини r та r' за формулами

$$r = \sum_{i=1}^{n} x_i{}^{C} \bmod q, \tag{30}$$

$$r' = \sum_{i=1}^{n} x_i{}^{E} \bmod q, \tag{31}$$

де $x^C = (x_1{}^C, x_2{}^C, ..., x_n{}^C)$, $x^E = (x_1{}^E, x_2{}^E, ..., x_n{}^E)$.

Абонент Б засліплює хеш-образ повідомлення за співвідношенням

$$h' = \frac{r'}{r} \cdot h \cdot \alpha \bmod q \tag{32}$$

і пересилає засліплений хеш-образ повідомлення h' підписувачу А.

Підписувач А формує засліплений підпис s' за допомогою власного секретного ключа d

$$s' = (d \cdot r' + k \cdot h') \bmod q, \tag{33}$$

та пересилає отримане значення абоненту Б.

Абонент Б має можливість перевірити справжність засліпленого підпису s' за допомогою співвідношення (34), використовуючи відкритий ключ Q підписувача А

$$s' \cdot P = r' \cdot Q + h' \cdot E. \tag{34}$$

Якщо засліплений підпис проходить перевірку, абонент Б формує з нього остаточний підпис

$$s = (s' \cdot \frac{r}{r'} + \beta \cdot h) \bmod q. \tag{35}$$

Сліпим підписом під документом m вважається пара значень $< r, s >$.

Валідатор при перевірці підпису $\{m, < r, s >\}$ обчислює точку R, використовуючи відкритий ключ Q підписувача А

$$R = \frac{s}{h} \cdot P - \frac{r}{h} \cdot Q = (x^R, y^R). \tag{36}$$

Підпис вважається справжнім, якщо виконується співвідношення

$$r = \sum_{i=1}^{n} x_i{}^{R} \bmod t. \tag{37}$$

де $x^R = (x_1{}^R, x_2{}^R, ..., x_n{}^R)$.

4 ОБЧИСЛЮВАЛЬНИЙ ПРИКЛАД ПРОТОКОЛУ СЛІПОГО ПІДПИСУ НА ЕК НАД СВП

Побудуємо СВП на основі простого поля $GF(11)$ з довжиною вектора $n = 2$. Правило множення задамо за табл. 1. З огляду на систему (5) значення розтягуючого коефіцієнта τ необхідно обирати з набору чисел $\{2, 6, 7, 8, 10\}$, які є квадратичними незалишками за модулем 11. Оберемо $\tau = 7$.

Розглянемо ЕК з коефіцієнтами $a = (1;3)$, $b = (5;6)$. Перевіримо умову гладкості (12) $4 \cdot (1;3) \circ (1;3) \circ (1;3) + 27 \cdot (5;6) \circ (5;6) = (0;3) \neq (0,0)$.

Базова точка $P = ((3;8),(4;9))$ має простий порядок $t = 113$.

Згенеруємо ключі підписувача А. Оберемо секретний ключ $d = 56$ та отримаємо за формулою (26) відкритий ключ $Q = 56 \cdot P = ((9;3),(9;9))$.

Підписувач А обирає одноразовий ключ $k = 28$ та обчислює за формулою (27) точку $E = 28 \cdot P = ((7;4),(0;3))$.

Абонент Б емітує документ з хеш-образом $h = 100$, обирає маскуючі параметри $\alpha = 44$ та $\beta = 75$, обчислює за формулою (29) точку $C = 44 \cdot E + 75 \cdot P = ((8;5),(10;0))$.

Після цього абонент Б обчислює за формулами (30–31) $r = x_1{}^C + x_2{}^C = 8 + 5 = 13$ та $r' = x_1{}^E + x_2{}^E = 7 + 4 = 11$ відповідно.

Абонент Б за формулою (32) засліплює повідомлення $h' = \frac{11}{13} \cdot 100 \cdot 44 \bmod 113 = 81$ та пересилає отримане значення підписувачеві А.

Підписувач А обчислює за формулою (33) засліплений підпис для отриманого значення $s' = (11 \cdot 56 + 28 \cdot 81) \bmod 113 = 59$ та відправляє його абонентові Б.

Абонент Б перевіряє справжність підпису за співвідношенням (34): в лівій частині співвідношення він отримує $59 \cdot P = ((5;2),(2;5))$, а в правій $11 \cdot Q + 81 \cdot E = ((5;2),(2;5))$. Оскільки значення в правій та лівій частинах перевірочного співвідношення збігаються, засліплений підпис вважається справжнім, і абонент Б обчислює за формулою (35) остаточний підпис $s = (59 \cdot \frac{13}{11} + 75 \cdot 100) \bmod 113 = 9$.

В результаті, під документом з хеш-образом $h = 100$ сформовано наосліп підпис $< r, s > = < 13, 9 >$.

Виконаємо перевірку сформованого підпису. За формулою (36) обчислимо точку $R = \frac{9}{100} \cdot P - \frac{13}{100} \cdot Q = ((8;5),(10;0))$. В правій частині перевірочного співвідношення (37) отримаємо $x_1{}^R + x_2{}^R = 8 + 5 = 13$. Ця величина збігається з $r = 13$, відповідно, підпис вважається справжнім.

Варто відзначити, що використання СВП в схемі сліпого ЕЦП не впливає на розмір підпису, зокрема, не призводить до його збільшення. Також важливо, що в процесі постановки підпису підписувач А не може дізнатися ні оригінального хеш-образу документу h, ні остаточного підпису $< r, s >$ під ним.

5 ПЕРЕВІРКА ЗАХИЩЕНОСТІ ПРОТОКОЛУ ЗА КРИТЕРІЄМ АНОНІМНОСТІ

Для схем сліпого підпису, на відміну від інших різновидів ЕЦП, актуальною є атака порушення анонімності. Спроба атаки може бути здійснена підписувачем за умови, що він зберігатиме всі відомі йому параметри схеми сліпого підпису разом із ідентифікатором емітента для кожної сесії постановки підпису. Накопичена база даних може бути використана в атаці, яка полягає у спробі визначення автора відомого документа m із підписом $< r, s >$, що проходить перевірку за допомогою відкритого ключа підписувача Q.

В запропонованому протоколі атака порушення анонімності може бути здійснена наступним чином. Підписувач А для кожного рядка своєї бази даних має обчислити ймовірні засліплюючи параметри α' та β' за формулами (38, 39)

$$\alpha' = \frac{r \cdot h'}{h \cdot r'} \bmod t, \tag{38}$$

$$\beta' = \frac{s - s' \cdot \dfrac{r}{r'}}{h} \bmod t, \tag{39}$$

За допомогою обчислених параметрів підписувач А для кожного рядка бази даних обчислює точку R' за формулою (40)

$$R' = \alpha' \cdot E + \beta' \cdot P = (x^{R'}, y^{R'}). \tag{40}$$

Рядок бази даних, для якого виконається співвідношення (41) має вказати на емітента повідомлення

$$r = \sum_{i=1}^{n} x_i^{R'} \bmod t. \tag{41}$$

Як доведено автором в [4], точка R' не залежить від параметрів h', r', s' і завжди збігається з перевірочною точкою R (див. співвідношення (36)), що не дає підписувачеві можливості визначити емітента.

$$R' = \alpha' \cdot E + \beta' \cdot P = \frac{r \cdot h'}{h \cdot r'} \cdot E + \frac{s - s' \cdot \dfrac{r}{r'}}{h} \cdot P =$$

$$= \frac{r}{h \cdot r'} \cdot \left(s' \cdot P - r' \cdot Q \right) + \frac{s}{h} \cdot P - \frac{s' \cdot r}{h \cdot r'} \cdot P = \frac{s}{h} \cdot P - \frac{r}{h} \cdot Q = R.$$

Таким чином, розглянутий протокол вважається захищеним за критерієм анонімності.

ВИСНОВКИ

В статті розглядається протокол сліпого електронного підпису на основі стандарту ГОСТ Р 34.10-2001 з використанням математичного апарату еліптичних кривих над скінченим векторним полем. Показано, що зміна математичного апарату не впливає на захищеність схеми сліпого підпису за критерієм анонімності.

В подальшому автором планується оцінка виграшу в швидкодії та ресурсомісткості при використанні математичного апарату ЕК над СВП в схемах ЕЦП з рекомендованими параметрами.

СПИСОК ЛІТЕРАТУРИ

1. *Молдовян, Н. А.* Теоретический минимум и алгоритмы цифровой подписи / Н. А. Молдовян. – С. Пб. : БХВ-Петербург, 2010. – 304 с.
2. Алгоритмические основы эллиптической криптографии / [Болотов А. А., Гашков С. Б., Фролов А. Б., Часовских А. А.]. – М. : Мэи, 2000. – 100 с.
3. *Костин, А. А.* О реализации протоколов слепой подписи и коллективной подписи на основе стандартов цифровой подписи / Костин А. А., Молдовян Н. А.,Фаль А. М. // Материалы VI Санкт-Петербургской межрегиональной конференции «Информационная безопасность России (ИБРР-2009). Санкт-Петербург, 28 30 октября 2009. – С. Пб. : СПОИСУ, 2009. – С. 111.
4. *Нікуліщев, Г. І.* Анонімність як критерій оцінки захищеності протоколів сліпого електронного цифрового підпису / Г. І. Нікуліщев, Г. Л. Козина // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. – 2012. – № 2.– С. 52–59.

Стаття надійшла до редакції 19.09.2013.

Никулищев Г. И.
Старший преподаватель, Запорожский национальный технический университет, Украина

ПРОТОКОЛ СЛЕПОЙ ЭЛЕКТРОННОЙ ЦИФРОВОЙ ПОДПИСИ НА ЭЛЛИПТИЧЕСКИХ КРИВЫХ НАД КОНЕЧНЫМ ВЕКТОРНЫМ ПОЛЕМ

В данной статье предлагается реализация протокола слепой электронной цифровой подписи, представляющего собой модификацию стандарта ГОСТ Р 34.10-2001 на эллиптических кривых над конечным векторным полем. Анализируется защищенность предлагаемого протокола по критерию анонимности.

Ключевые слова: слепая электронная цифровая подпись, конечное векторное поле, эллиптическая кривая.

Nikulishchev H. I.
Senior tutor, Zaporizhzhia national technical university, Ukraine

BLIND DIGITAL SIGNATURE PROTOCOL ON ELLIPTIC CURVES OVER VECTOR FINITE FIELD

Digital signature schemes can fulfill an actual task of ensuring fairness and transparency of electronic election. However, existing standards and protocols need modification into blind signature and additional security check by the anonymity criterion. The author examines blind signature protocol provided by Russian scientists. It is proposed to improve scheme's efficiency by changing inner mathematics. Elliptic curves over vector finite field enable parallel processing in group operation and reduce integer range. These advantages are illustrated by computational example. Also, improved protocol investigation by the anonymity criterion is provided in the article. The author proves that mathematics change do not affect protocol security.

Keywords: blind digital signature, vector finite field, elliptic curve.

REFERENCES

1. Moldovyan N. A. Teoreticheskij minimum i algoritmy' cifrovoj podpisi. Sankt-Peterburg, BXV-Peterburg, 2010, 304 p.
2. Bolotov A. A., Gashkov S. B., Frolov A. B., Chasovskix A. A. Algoritmicheskie osnovy' e'llipticheskoj kriptografii. Moscow, Me'i, 2000, 100 p.
3. Kostin A. A., Moldovyan N. A., Fal' A. M. O realizacii protokolov slepoj podpisi i kollektivnoj podpisi na osnove standartov cifrovoj podpisi. *Materialy' VI Sankt-Peterburgskoj mezhregional'noj konferencii «Informacionnaya bezopasnost' Rossii (IBRR-2009)». Sankt-Peterburg, 28–30 oktyabrya 2009,* Sankt-Petersburg, SPOISU, 2009, pp. 111.
4. Nikulishchev H. I., Kozina H. L. Anonimnist yak kryterii otsinky zakhyshchenosti protokoliv slipoho elektronnoho tsyfrovoho pidpysu. *Pravove, normatyvne ta metrolohichne zabezpechennia systemy zakhystu informatsii v Ukraini,* 2012, No. 2, pp. 52–59.

УДК 517.9

Панасенко Є. В.

Канд. фіз.-мат. наук, ст. викладач, Запорізький національний університет, Україна, E-mail: panasenko.yevgeniy@gmail.com

РОЗВ'ЯЗОК КРАЙОВОЇ ЗАДАЧІ ОПТИМАЛЬНОГО КЕРУВАННЯ В КРИТИЧНОМУ ВИПАДКУ

У статті розглянуто крайову задачу в критичному випадку, яка виникає в теорії оптимального керування для матричних диференціальних рівнянь. Знайдено умову розв'язності таких задач. Запропоновано підхід до знаходження її розв'язку за допомогою теорії псевдообернених матриць.

Ключові слова: керуючий процес, крайова задача, псевдообернена матриця, нормальна фундаментальна матриця.

ПОСТАНОВКА ПРОБЛЕМИ

Всюди в навколишньому світі протікають різні процеси, характер яких залежить від багатьох умов і факторів. Змінюючи умови протікання процесів, людина може впливати на їх характер, змінювати їх, пристосовувати до своїх цілей. Це втручання в природний хід процесу і являє собою сутність керування в широкому сенсі слова. Можна сказати, що керування представляє собою таку організацію того чи іншого процесу, яке забезпечує досягнення певних цілей.

При розгляданні реальних керованих об'єктів перш за все виникає задача керування рухом [1], яка зводиться до розв'язку крайової задачі для системи звичайних диференціальних рівнянь першого порядку, розмірності $n \times n$. Розв'язання таких задач приводить до певних труднощів, що характерно для випадків керування системами, в яких кількість процесів перевищує кількість відомостей про початковий стан. На практиці розв'язність таких прикладних задач ґрунтується на «фізичних міркуваннях» та немає строгого теоретичного підґрунтя.

Важливим частинним випадком є лінійні звичайні системи матричних диференціальних рівнянь, які моделюють лінійні технологічні та економічні процеси [2]. Це системи наступного вигляду:

$$\frac{dx}{dt} = A(t)x(t) + B(t)u, \quad t_0 \le t \le T, \tag{1}$$

де $x(t) \in \mathrm{R}^n$ – фазовий вектор, $u(t) \in \mathrm{R}^m$ – вектор керування. Матриці $A(t)$ і $B(t)$ припускаються неперервними, розмірності $n \times n$ і $n \times m$ відповідно. Допустимим керуванням вважається довільне кусково-неперервне керування $u = u(t)$, причому всі точки розриву функції $u = u(t)$ – першого роду (якщо такі є). Кожному такому допустимому керуванню відповідає єдиний розв'язок крайової задачі:

$$\frac{dx}{dt} = A(t)x(t) + B(t)u(t), \quad t_0 \le t \le T, \tag{2}$$

$$Mx(t_0) = \alpha, \tag{3}$$

де M – матриця розмірності $m \times n$, $\alpha \in \mathrm{R}^m$. Зокрема, цікавлять умови існування розв'язку крайової задачі (2)–(3) і визначення функції $x(t)$, яка задовольняє крайовій умові (3) при відомому вхідному процесі $u(t)$. Якщо процес $x(t)$ отриманий, то визначення виходу системи $y(t)$ не представляється складним і виконується безпосередньо по рівнянню виходу $y(t) = C(t)x(t) + D(t)u(t)$ [3, 4].

АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Сьогодні у теорії керування при аналізі лінійних систем традиційно використовують підходи, які є досить відомими в теорії диференціальних рівнянь [5, 6, 7].

В залежності від вигляду крайової умови (3) отримуємо різні крайові задачі. Наприклад, в роботі [3] побудовано розв'язок для задачі Коші, коли $M = I_n$, де I_n – одинична матриця (визначено функцію $x(t)$ по заданому початковому стану x_0 при відомому вхідному процесі $u(t)$), використовуючи поняття фундаментальної (перехідної або імпульсної) матриці [2, 3] у вигляді

$$x(t) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)d\tau, \tag{4}$$

де $\Phi(t, t_0) = X(t)X^{-1}(t_0)$.

В роботі [4] розглянута крайова задача при $t \in [0; T]$ і крайовою умовою загального вигляду $\ell x(\cdot) = \alpha$, $\alpha \in \mathrm{R}^n$, де $\ell : D^n[0; T] \to \mathrm{R}^n$ – лінійний обмежений вектор-функціонал. Для розв'язності задачі вимагається, щоб

$$\det[\ell X] \ne 0. \tag{5}$$

Таким чином розв'язність крайової задачі не залежить від неоднорідностей диференціального рівняння (2) і від $\alpha \in \mathrm{R}^n$. Із вигляду (5), очевидно, що розв'язність задачі залежить тільки від фундаментальної матриці $X(t)$ та вектор-функціонала ℓ. При виконанні умови (5) розв'язок

задачі дається через матрицю Гріна

$$x(t) = X(t)[\ell X]^{-1}\alpha + \int_0^T G(t,\tau)B(\tau)u(\tau)d\tau . \qquad (6)$$

В даній роботі розглядається випадок, коли не вимагається виконання умови (5). Більш того, не вимагається, навіть, щоб матриця $[\ell X]$ була квадратною. В цьому випадку отримаємо крайову задачу, в якій кількість крайових умов не співпадає з кількістю невідомих у системі [6, 7].

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Наряду з неоднорідною крайовою задачею (2)–(3), розглянемо однорідну крайову задачу

$$\frac{dx}{dt} = A(t)x(t), \quad t_0 \le t \le T, \qquad (7)$$

$$Mx(t_0) = 0 . \qquad (8)$$

Розв'язок $x(t)$ рівняння (2)

$$x(t) = x_0 + \int_{t_0}^{t} (A(s)x(s) + B(s)u(s))ds, \qquad (9)$$

неперервно-диференційований в кожній точці $t \in [t_0;T]$ і задовольняє рівнянню (2) всюди на відрізку $[t_0;T]$. Отже розв'язок $x(t)$ рівняння (2) будемо шукати в просторі $C^1([t_0;T])$ неперервно-диференційованих на $[t_0;T]$ функцій.

Знайдемо умови існування та структуру розв'язків неоднорідної скінченновимірної крайової задачі (2), (3).

Використовуючи апарат теорії псевдообернених матриць [6, 7], знайдемо необхідні та достатні умови розв'язності крайової задачі (2), (3) в просторі неперервно-диференційованих на відрізку $[t_0;T]$ функцій. Позначимо через $X(t,\tau)$ – нормальну фундаментальну матрицю, яка задовольняє матричному рівнянню:

$$\frac{dX(t,\tau)}{dt} = A(t)X(t,\tau);$$

$X(\tau,\tau) = I$ – одинична матриця. Загальний розв'язок матричного рівняння має вигляд [8]

$$x(t) = X(t)x_0 + \bar{x}(t), \qquad X(t) = X(t,t_0), \qquad (10)$$

де $x_0 = x(t_0)$ – елемент простору R^n, $\bar{x}(t)$ – частинний розв'язок неоднорідного рівняння (2), який може бути записаний у вигляді

$$\bar{x}(t) = X(t)\int_{t_0}^{t} X^{-1}(\tau)B(\tau)u(\tau)d\tau.$$

Підставимо (10) в крайову умову (3) та отримаємо наступне алгебраїчне рівняння відносно елемента $x_0 \in \mathrm{R}^n$:

$$Qx_0 = \alpha, \qquad (11)$$

де $Q = MX(t_0)$ – матриця, отримана підстановкою в крайову умову (3) нормальної фундаментальної матриці $X(t)$ однорідної системи (7).

Відомо [5], що алгебраїчна система (11) розв'язна тоді і тільки тоді, коли права частина належить ортогональному доповненню $^{\perp}N(Q^*) = R(Q)$ підпростору $N(Q^*)$, тобто виконується умова

$$\mathrm{P}_{N(Q^*)}\alpha = 0. \qquad (12)$$

При цьому загальний розв'язок системи (11) має вигляд

$$x_0 = Q^+\alpha + \mathrm{P}_{N(Q)}c, \qquad (13)$$

де Q^+ – псевдообернена матриця по Пенроузу [9] до матриці Q; c – довільний елемент з простору R^n.

Підставимо x_0 у вираз (10), отримаємо загальний розв'язок крайової задачі (2), (3) у вигляді

$$x(t) = X(t)\mathrm{P}_{N(Q)}c + X(t)Q^+\alpha + K[f](t), \qquad (14)$$

де $K[f](t)$ – оператор Гріна задачі (2), (3), який діє на функцію $f(t) \in C([t_0;T])$ наступним чином:

$$K[f](t) := X(t)\int_{t_0}^{t} X^{-1}(\tau)B(\tau)u(\tau)d\tau . \qquad (15)$$

Випадки крайових задач, для яких виконана одна з умов $\operatorname{rank} Q = n$ або $\operatorname{rank} Q < n$, є відповідно некритичними і критичними, де $Q = MX(t_0)$ – $(m \times n)$ – вимірна матриця, отримана підстановкою у крайову умову нормальної фундаментальної матриці $X(t)$ однорідної системи (7).

Таким чином, доведено наступне твердження.

Теорема. Якщо $\operatorname{rank} Q = n_1$, то однорідна крайова задача (7)–(8) має $r = n - n_1$ і тільки r лінійно незалежних розв'язків. Неоднорідна крайова задача (2)–(3) розв'язна тоді і тільки тоді, коли виконана умова $\mathrm{P}_{N(Q^*)}\alpha = 0$ і при цьому має r-параметричну родину розв'язків

$$x(t,c_r) = X(t)\mathrm{P}_{N(Q)}c_r + X(t)Q^+\alpha + \\ + X(t)\int_{t_0}^{t} X^{-1}(\tau)B(\tau)u(\tau)d\tau , \qquad (16)$$

де $\mathrm{P}_{N(Q)} = I - Q^+Q$, $\mathrm{P}_{N(Q^*)} = I - QQ^+$ – ортопроектори на ядро $N(Q)$ і коядро $N(Q^*)$ матриці Q відповідно, Q^+ – псевдообернена матриця до матриці Q.

Приклад

Нехай керуючий процес описується крайовою задачею (2)–(3), де $x(t) \in \mathrm{R}^3$ – фазовий вектор, $t_0 = \sqrt{3} \le t \le 3 = T$, $u(t) = \left(1 + t^2\right) \in \mathrm{R}^1$ – вектор керування та задані матриці $A(t) = \begin{pmatrix} \frac{2t}{1+t^2} & 0 & 0 \\ 0 & \frac{2t}{1+t^2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$,

$B(t) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $M = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -1 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix}$, $\alpha = \begin{pmatrix} 18 \\ 9 \end{pmatrix}$.

Нормальна фундаментальна матриця для заданої задачі має вигляд $X(t) = \begin{pmatrix} 1+t^2 & 0 & 0 \\ 0 & 1+t^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Рівняння (11) представляє собою систему алгебраїчних рівнянь

$$\begin{pmatrix} 2 & -2 & -1 \\ 0 & -2 & 1 \end{pmatrix} x_0 = \begin{pmatrix} 18 \\ 9 \end{pmatrix} \qquad (17)$$

відносно елемента $x_0 \in \mathrm{R}^3$. $\mathrm{P}_{N(Q^*)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, тому умова розв'язності (12) для даної крайової задачі виконана насправді для довільних $\alpha \in \mathrm{R}^2$ у тому числі і для $\alpha = \begin{pmatrix} 18 \\ 9 \end{pmatrix} \in \mathrm{R}^2$. Система (17) є недовизначеною, розв'язок якої знаходиться за формулою (13), де

$$Q^+ = \begin{pmatrix} \frac{5}{18} & -\frac{1}{6} \\ -\frac{1}{9} & -\frac{1}{3} \\ -\frac{2}{9} & \frac{1}{3} \end{pmatrix}, \quad \mathrm{P}_{N(Q)} = \begin{pmatrix} \frac{4}{9} & \frac{2}{9} & \frac{4}{9} \\ \frac{2}{9} & \frac{1}{9} & \frac{2}{9} \\ \frac{4}{9} & \frac{2}{9} & \frac{4}{9} \end{pmatrix}.$$

Неоднорідна крайова задача має 3-параметричну родину розв'язків (формула (16)):

$$x(t) = \begin{pmatrix} \left(1+t^2\right)\left(t-\sqrt{3}\right) + \frac{7}{2} + \frac{7}{2}t^2 + \left(\frac{4}{9} + \frac{4}{9}t^2\right)c_1 + \left(\frac{2}{9} + \frac{2}{9}t^2\right)c_2 + \left(\frac{4}{9} + \frac{4}{9}t^2\right)c_3 \\ \left(1+t^2\right)\left(t-\sqrt{3}\right) - 5 - 5t^2 + \left(\frac{2}{9} + \frac{2}{9}t^2\right)c_1 + \left(\frac{1}{9} + \frac{1}{9}t^2\right)c_2 + \left(\frac{2}{9} + \frac{2}{9}t^2\right)c_3 \\ -1 + \frac{4}{9}c_1 + \frac{2}{9}c_2 + \frac{4}{9}c_3 \end{pmatrix} \in \mathrm{R}^3. \qquad (18)$$

У достовірності отриманого результату можна переконатися елементарною перевіркою, а саме, підставити (18) в крайову задачу (2)–(3).

ВИСНОВКИ

Робота присвячена знаходженню умов розв'язності та побудови розв'язку критичних крайових задач, які виникають в задачах теорії керування руху, в яких кількість крайових умов не співпадає з кількістю невідомих у системі. Крім того, даний підхід можна застосувати до задач про аналітичне конструювання регуляторів і про оптимальну стабілізацію [2, 6], при розв'язку яких виникають матричні рівняння Ріккаті.

СПИСОК ЛІТЕРАТУРИ

1. *Ванько, В. И.* Вариационное исчисление и оптимальное управление: Учеб. для вузов / В. И. Ванько, О. В. Ермошина, Г. Н. Кувыркин. – М. : Изд-во МГТУ им. Н. Э. Баумана, 2006. – 488 с.
2. *Егоров, А. И.* Обыкновенные дифференциальные уравнения с приложениями. 2-е изд., испр. / А. И. Егоров. – М. : ФИЗМАТЛИТ, 2005. – 384 с.
3. *Андриевский, Б. Р.* Избранные главы теории автоматического управления с примерами на языке MATLAB / Б. Р. Андриевский, А. Л. Фрадков. – С.Пб. : Наука, 2000. – 475 с.
4. *Максимов, В. П.* Теория оптимального управления. Часть 2 элементы теории линейных операторов и операторных уравнений / В. П. Максимов, П. М. Симонов. – Пермь : Перм. гос. ун-т, 2010. – 80 с.
5. *Бойчук, А. А.* Конструктивные методы анализа краевых задач / А. А. Бойчук. – К. : Наукова думка, 1990. – 96 с.
6. *Бойчук, А. А.* Обобщенно-обратные операторы и нетеровы краевые задачи / А. А. Бойчук, В. Ф. Журавлев, А. М. Самойленко. – К. : Институт математики НАНУ, 1995. – 320 с.
7. *Boichuk, A. A.* Generalized inverse operators and fredholm boundary-value problems / A. A. Boichuk, A. M. Samoilenko. – VSP, Utrecht-Boston, 2004. – 317 p.
8. *Далецкий, Ю. Л.* Устойчивость решений дифференциальных уравнений в банаховом пространстве / Ю. Л. Далецкий, М. Г. Крейн. – М. : Наука, 1970. – 536 с.
9. *Penrose, R.* Generalized Inverse for Matrices / R. Penrose // Proc. Cambriadge Philos. Soc. – 1955. – Vol. 51, № 3. – P. 406–413.

Стаття надійшла до редакції 14.03.2013.

Панасенко Е. В.
Канд. физ.-мат. наук, доцент, Запорожский национальный университет, Украина

РЕШЕНИЕ КРАЕВОЙ ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ В КРИТИЧЕСКОМ СЛУЧАЕ

В статье рассмотрена краевая задача в критическом случае, которая возникает в теории оптимального управления для матричных дифференциальных уравнений. Найдено условие разрешимости таких задач. Предложен подход к нахождению решения задачи с помощью теории псевдообратных матриц. Такие задачи могут возникать при моделировании линейных технологических и экономических процессов. Кроме того, данный подход применим к задачам об аналитическом конструировании регуляторов и об оптимальной стабилизации.

Ключевые слова: управляемый процесс, краевая задача, псевдодбратная матрица, нормальная фундаментальная матрица.

Panasenko Y. V.
Ph.D. Candidate (Candidate of Physico-mathematical Sciences), Associate Professor, Zaporizhzhya National University, Ukraine

SOLUTION OF THE BOUNDARY-VALUE PROBLEM OF OPTIMAL CONTROL IN THE CRITICAL CASE

The article describes the boundary-value problem of the theory of optimal control for matrix differential equations in the critical case. This problem is urgent for wide range of applications like modeling of linear technological and economic processes. A problem of controlling movement is reduced to solutions of boundary-value problems of $n \times n$ systems of ordinary differential equations of the first order. Solution of these tasks is very complex in case of control of systems in which the number of processes exceeds the amount of information about the initial state.

The paper is devoted to finding conditions for solvability and construction solutions of boundary-value problems in critical case of the theory of motion control, in which the number of boundary conditions does not coincide with the number of unknowns in the system. The condition of solvability of such problems is found. Author describes the approach for solution to the problem with the help of theory of pseudoinverse matrices.

In addition this approach is applicable to the problem of analytic construction of regulators and the optimal stabilization.

Keywords: controlled process, boundary-value problem, pseudoinverse matrix, normal fundamental matrix.

REFERENCES

1. Vanko V. I., Yermoshina O. V., Kuvyrkin G. N. Variatsionnoye ischisleniye i optimalnoye upravleniye: Ucheb. dlya vuzov. Moscow, Izd-vo MGTU im. N. E. Baumana, 2006, 488 p.
2. Egorov A. I. Obyknovennye differentsialnye uravneniya s prilozheniyami. 2-e izd., ispr. Moscow, FIZMATLIT, 2005, 384 p.
3. Andrievskii B. R., Fradkov A. L. Izbrannye glavy teorii avtomaticheskogo upravleniya s primerami na yazyke MATLAB, Sankt-Peterburg, Nauka, 2000, 475 p.
4. Maksimov V. P., Simonov P. M. Teoriya optimalnogo upravleniya. Chast 2 elementy teorii lineinykh operatorov i operatornykh uravnenii. Perm, Perm. gos. un-t, 2010, 80 p.
5. Boichuk A. A. Konstruktivnye metody analiza kraevykh zadach. Kiev, Naukova dumka, 1990, 96 p.
6. Boichuk A. A., Zhuravlev V. F., Samoilenko A. M. Obobshchenno-obratnye operatory i neterovy kraevye zadachi. Kiev, Institut matematiki NANU, 1995, 320 p.
7. Boichuk A. A., Samoilenko A. M. Generalized inverse operators and fredholm boundary-value problems. VSP, Utrecht-Boston, 2004, 317 p.
8. Daletskii Yu. L., Krein M. G. Ustoichivost reshenii differentsialnykh uravnenii v banakhovom prostranstve. Moscow, Nauka, 1970, 536 p.
9. Penrose R. Generalized Inverse for Matrices, *Proc. Cambriadge Philos. Soc.*, 1955, Vol. 51, No. 3, pp. 406–413.

UDC 519.766.4

Trofymchuk O. M. [1], Kozhukhivska O. A. [2], Bidyuk P. I. [3], Kozhukhivskyi A. D. [4]

[1] *Doctor of technical sciences, professor, deputy director of Institute of telecommunications and information technologies of NAS Ukraine, Ukraine*

[2] *Candidate of technical sciences, senior teacher, Cherkassy state technological university, Ukraine*

[3] *Doctor of technical sciences, professor, Institute of applied system analysis of National technical university of Ukraine «KPI», Ukraine*

[4] *Doctor of technical sciences, professor of Cherkassy state technological university, Ukraine E-mail: andrejdk@mail.ru*

ESTIMATION OF MARKET RISK IN UKRAINE USING VAR METHODOLOGY

The emergence of a market risk due to performing operations with currency can result in substantial financial losses. That is why such situations require carrying out of profound analysis and management of respective risks. The market risk of this kind is characterized with possible losses of financial resources due to incorrectly performed operations with currency. The paper considers the possibility of application of the *VaR* methodology to the bank currency portfolio using the following methods: delta-normal, as well as the methods of historical modeling and Monte-Carlo simulation. While performing the computing experiments actual data used from the currency market of Ukraine. Quite acceptable results of forecasting possible losses were received by making use of Monte-Carlo simulation that hypothetically can take into account possible variations of the market exchange rates. It was established that the risk forecasting errors appear only due to non-predictable abrupt changes of exchange rates.

Keywords: risk measure VaR, bank currency portfolio, historical modeling, Monte-Carlo simulation, delta-normal approach, Ukrainian currency market.

1 INTRODUCTION

As far as functioning of financial institutions is closely related to performing of substantial volumes of currency operations the problem naturally arises for performing profound analysis and management of possible financial risks. The currency related risk is a possibility of financial resources loss due to incorrectly performed operations with currency. From the risk management position the banking activities are basically directed towards risk acceptance and getting respective economic compensation instead. Some types of risks represent the price of banking business realization and it is impossible to avoid them completely. That is why the risk management processes is not aiming to complete elimination of the risks. A financial institution should provide a reliable substantiated relation between generalized parameters of possible risks and the capital, available financial resources and financial incomes [1, 2].

There exist various approaches to quantitative estimation of possible losses. As of today there are developed the methods for computing of the currency risks that are widely used in financial enterprises. A selection of appropriate computing method is determined by the volume of available information, qualification of personnel that is busy with risk management problems, and the availability of necessary working instrumentation in the form of computer software. In spite of the fact that such instrumentation market for the financial analysis includes rather wide choice possibilities their cost and the practical usage problems very often result in development of their own software products by the financial institutions. Such systems for risk estimation may exhibit much more functional restrictions that availa than available at the market but their advantages are in the possibility of fast extension of a number of practically needed functions. Also in such cases the financial institution personnel has a possibility to enhance substantially their qualification and to improve existing computing methodologies.

The paper is devoted to application of the *VaR* methodology for computing possible financial losses in analysis of a currency market with the use of original software.

2 THE PROBLEM STATEMENT

The goal of the work is to execute the analysis of influence of exchange rates oscillations on profitability of currency transactions; to present algorithms of calculation of VaR meanings using delta–normal method and also using methods of historical and imitation modeling; to make comparison analysis of using indicated methods of VaR estimation and give recommendations concerning possibilities of their usage on Ukrainian financial market.

3 THE INFLUENCE OF EXCHANGE RATES OSCILLATIONS ON PROFITABILITY

The model of currency matching. Despite the fact that all financial risks in this or that way are implemented on the results of bank activity but the functional connection between risks exists not for all its types. The dependence between size of profits (losses) received as a result of bank holding an open currency line item, and market changes of currency rates is described by a model of currency matching [3]: $\Delta P_v = VP(s_p - s)$, where ΔP_v is a profit (loss) received from overestimating of currency money because of currency rate change; VP is a currency line item of a bank; s_p, s – predicted and leaking currency rates accordingly.

A currency item (CI) is the indicator of bank's currency risk that is defined by the parity between the amount of assets in certain foreign currency (A_v) and the amount of obligations in that same currency (L_v): $VP = A_v - L_V$. CI of a bank can be open or closed and be calculated separately for each foreign currency that is included in multicurrency bank briefcase. CI is considered to be open if the amount of assets in foreign currency does not correspond the amount of obligations in that same currency. If the amount of assets in foreign currency is balanced by the amount of obligations in that same foreign currency ($A_v = L_v$), such item is called closed or the item of correspondence. In such case currency risk is almost absent because the rate change of one currency concerning the other will equally influence both the cost of assets and the cost of obligations.

4 THE ESTIMATION OF CURRENCY RISK VAR USING DELTA-NORMAL METHOD

In order to demonstrate shortcomings and advantages of a delta-normal method, let's consider how to estimate possible future changes of the cost of briefcase of currency money.

The algorithm of calculation of VaR. The cost of briefcase of currency money P_t in base currency is calculated with an expression:

$$P_t = \sum_{i=1}^{N} P_t^i = \sum_{i=1}^{N} k_t^i \cdot v_t^i,$$

where P_t – time line of costs of the whole briefcase of currency money in base currency in the moment of time t ($t = \overline{0,T}$, where $T + 1$ – quantity of meanings of time line P_t); $P_t^i = k_t^i \cdot v_t^i$ – the cost of briefcase component in i currency in base currency; k_t^i – exchange rate of briefcase i currency into base currency on date t ($i = \overline{1,N}$, where N – quantity of currency in briefcase); v_i^t – the volume of i currency in briefcase on date t (the size of open currency item in terms of currencies). Let's consider the sequence of calculations of risk cost VaR that reflects the possible volume of future changes of the cost of currency money briefcase P_t.

Stage 1. The calculation of daily change of currency rates. The meaning of daily change of rates of briefcase currencies is calculated with a formula of geometrical profitability:

$$x_t^i = \ln\left(\frac{k_t^i}{k_{t-1}^i}\right),$$

where k_t^i – the meaning of exchange rate of i currency to base currency on date t, $t = \overline{1,T}$; k_{t-1}^i – the meaning of exchange rate of i currency to base currency on date t *1*. Logarithm of time of changes of currency rate characterizes the intensity of change of currency rate and is a random variable, the distributing of which is close to normal with average meaning close to zero.

Stage 2. The calculation of volatility of currencies. In order to calculate the volatility of each currency separately without taking into account its connection with other currencies in briefcase, it is necessary to calculate for each currency selective average and standard quadratic deviation σ^i time line of its profitability $\{x_t^i\}$ with an expression:

$$\sigma^i = \sqrt{\frac{\sum_{i=1}^{T}(x_t^i - \overline{X^i})^2}{T-1}}. \qquad (1)$$

Stage 3. The estimation of possible losses behind the open currency item in *i* currency VaR_i. The variable of risk cost VaR^i of open item in i currency is calculated with an expression:

$$VaR^i = k_{1-\alpha} P_t^i \sigma^i. \qquad (2)$$

If volatility of i currency (1) is defined on daily interval, the risk cost of VaR^i is also interpreted as maximum expected volume of reduction of total cost of a separately taken component of currency briefcase in i currency during one day with possibility 95 % or 99 % depending on the meaning of quantile $k_{1-\alpha}$ in the expression (2).

Stage 4. The calculation of correlation matrix of briefcase currencies. In order to consider mutual correlation of exchange rate of briefcase currencies in the process of VaR calculation it is necessary to find correlation matrix of briefcase currencies. To do this first it is necessary to calculate co variations C_{ij} of possiblecombinations of random variables $\{x_t^i\}$ and $\{x_t^j\}$:

$$C_{ij} = \frac{1}{T}\sum_{t=1}^{T}(x_t^i - \overline{X^i}) \cdot (x_t^j - \overline{X^j}).$$

And also correlation coefficients K_{ij} of processes $\{x_t^i\}$ and $\{x_t^j\}$: $K_{ij} = \dfrac{C_{ij}}{\sigma_i \sigma_j}$.

Square matrix with dimension $n \times n$, in which on i row and j column intersection the element K_{ij} is located, is a correlation matrix of briefcase exchange rates. This matrix is symmetric: $K_{ij} = K_{ji}$ for all i, $j = \overline{1,N}$, and the elements of main diagonal are single.

Stage 5. The calculation of overall estimation of possible losses VaR of total cost of currency briefcase. Overall estimation of possible losses of total cost of currency briefcase VaR is calculated on the basis of risk costs of VaR^i of separate currencies of a briefcase and correlation matrix of exchange currency rates:

$$VaR = \sqrt{\overline{VaR} \cdot \| \mathbf{K} \| \overline{VaR^T}},$$

where $\overline{VaR} = \left\| VaR^1 \quad VaR^2 \quad ... \quad VaR^N \right\|$ – vector-line of separate estimations VaR^i of briefcase parts in i currency; $\mathbf{K}$ – correlation matrix of exchange rates of briefcase currencies to base currency. This method assumes daily data updating and logarithm calculation of course growth rates, co variation and correlation matrixes, volatilities, all VaR^i estimations.

5 THE ESTIMATION OF CURRENCY BANK RISK VAR USING THE METHOD OF HISTORICAL MODELING

First it is necessary to choose the period of time with depth T (for example, 250 working days). For these days selection is created from daily changes of currency rates for all N parts of currency briefcase:

$$\Delta k_t^i = k_t^i - k_{t-1}^i, \; i = \overline{1,N},$$

where k_t^i – the meaning of exchange rate of i currency to base currency on date t, $t = \overline{1,T}$; k_{t-1}^i – the meaning of exchange rate of i currency to base currency on date t–1. For each of T scenarios of rate changes it is modeled hypothetical rate k^* of each currency in future as its current rate k_0 plus rate growth which corresponds the chosen scenario:

$$k_t^i{}^* = k_{i,0} + \Delta k_t^i.$$

Then it is conducted the complete revaluation of current currency briefcase according to rates modeled on the basis of historical scenarios, and for each scenario it is calculated how the cost of today (current) currency briefcase (separately according to long and short bank currency item) would change:

$$\Delta V_t = V_t^* - V_0, \; t = \overline{1,T},$$

where $V_0 = \sum_{i=1}^{N} k_{i,0} \cdot v_{i,0}$ – current cost of currency briefcase; $v_{i,0}$ – current volume of i currency in briefcase (the cost of open currency item in currency units); $V_t^* = \sum k_i^{t*} \cdot v_{i,0}$ – the cost of currency briefcase in base currency according to t historical scenario.

Received T changes of briefcase are ranged by falling for long currency item and on the contrary for the short. VaR is defined as maximum loss that is not exceeded in $(1-\alpha)T$ cases, or is equal to absolute variable of change with a number that corresponds integer part of a figure . This method is relatively easy to implement if daily updating data base of all currencies exists. As a rule the more depth of a retrospective that is used for rates modeling, the higher is the accuracy of estimation VaR, but at the same time the bigger is the risk of using out-of-date data.

6 THE METHOD OF IMITATION MODELING MONTE-CARLO FOR ESTIMATION OF BANK CURRENCY RISKS VAR

Monte-Carlo method for estimation of currency risk consists in modeling *movement path of exchange rate according to chosen stochastic process.* In order to calculate the estimation VaR^i cost of i part of currency briefcase (open currency item, $i = \overline{1,N}$) it is necessary to build the division of modeled costs for this part. To find the line of modeled future costs of i currency item it is necessary to model K future prices according to movement path which is calculated during n steps. The figures K and n are chosen quite big depending on calculating capacities (for example, 500*1000). The process of estimation VaR can be represented the way:

1. To generate the consequence according to divided random variables $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$.

2. Using retrospective data of depth L of days to find the meaning of daily pr ofitability (change of i rate of currencies) with a formula $x_l^i = \ln(\frac{k_l^i}{k_{l-1}^i})$, $l = \overline{1,L}$. For the received sampling of profitability for i currency to calculate average μ and mean squared deviation σ.

3. Starting with current rate of i currency k_t^i to calculate future modeled prices $k_{t+1}^i, k_{t+2}^i, ..., k_{t+n}^i = k_T^i$ with formulas [2]:

$$k_{t+1}^i = k_t^i + k_t^i(\mu \, \Delta t + \varepsilon_1 \sigma\sqrt{\Delta t}),$$

$$k_{t+2}^i = k_{t+1}^i + k_{t+1}^i(\mu \, \Delta t + \varepsilon_2 \sigma\sqrt{\Delta t}),$$

$$... \; ... \; ... \; ... \; ... \; ... \; ... \; ... \; ... \; ... \; ...$$

$$k_{t+n}^i = k_{t+n-1}^i + k_{t+n-1}^i(\mu \, \Delta t + \varepsilon_n \sigma\sqrt{\Delta t}).$$

4. To calculate the cost of i currency item in base currency (the part of currency briefcase) for rate k_T^i: $P_T^i = k_T^i \cdot v_t^i$, where v_t^i – current volume of i currency item in units of currencies.

5. The steps 3–4 to repeat K times (depending on the quantity of variables). As a result we get a line of meanings: $P_T^{i\,1}, P_T^{i\,2}, ..., P_T^{i\,K}$.

6. The received K costs of i currency item of briefcase are ranged similarly to the method of historical modeling.

The ranged meanings are numbered from 1 to K. Let's designate through $P_T^{i\,\alpha}$ the meaning of currency item from this line by a number that is equal to the whole part of a number $(1-\alpha)K$, that means that it corresponds the set level of trust $(1-\alpha)$.

7. To calculate the average of modeled costs:

$$\overline{P_T^i} = \frac{\sum_{k=1}^{K} P_T^{i\,k}}{K}.$$

8. To calculate possible losses according to i currency item: $VaR^i = \overline{P_T^i} - P_T^{i\,\alpha}$.

7 THE PECULIARITIES OF MODEL VERIFICATION TO ESTIMATE VAR CURRENCY BRIEFCASE USING HISTORICAL DATA

The further specified operations influence the size of an open currency position and currency risk: (1) – purchase and sale of available and non-cash foreign currency; (2) – charge, receiving, payment of foreign currency in a form of profits and losses; (3) – receipt of funds in foreign currency to statute capital; (4) – repayment by a bank of hopeless debt in foreign currency; (5) – formation of reserves in foreign currency; (6) – purchase and sale of inventory items using foreign currency; (7) – other exchange operations with foreign currency [3–7]. To estimate the changes of structure of bank currency item that does not depend on exchange rates oscillations, the index of Paashe is used.

Let's consider that the total cost of currency briefcase P_t in base currency on the moment of time t is defined by a formula:

$$P_t = \sum_{i=1}^{N} k_t^i \cdot \left| v_t^i \right|,$$

where $P_t^i = k_t^i \cdot v_t^i$ – currency item of a bank according to i currency in base currency. The index of Paashe J_{vi} characterizes the level of influence of structural changes of a currency item on the total volume of currency item taking into account exchange rates on the beginning of a period $[t-1,\ t]$ [3]:

$$J_{vi} = \frac{k_{t-1}^i \cdot v_t^i}{k_{t-1}^i \cdot v_{t-1}^i},\ i = \overline{1, N}. \qquad (3)$$

The process of verification of a model for currency risk estimation is the following. On the moment of time t it is possible to calculate the meaning of actual cost of a briefcase Δ_t and compare it with the meaning of VaR_t. The peculiarity of comparison consists of necessity to exclude from calculation the factor of change of physical volumes of briefcase currencies as far as the indicator VaR does not take into account the changes of volumes of each currency in briefcase. Taking this the adequacy check of VaR-model is made in the following sequence:

1. To define Δ_t^i as losses from i currency item for a period of time $[t-1, t]$, as disparity between the cost of i currency item on the moment of time t without taking into account the changes of physical structure that occurred on the period $[t-1, t]$ and its cost on the moment of time $t-1$ with a formula:

$$\Delta_t^i = \begin{cases} \left| \frac{P_t^i}{J_{vi}^t} - P_{t-1}^i \right|, \text{ if } \frac{P_t^i}{J_{vi}^t} - P_{t-1}^i < 0 \\ 0, \text{ if } \frac{P_t^i}{J_{vi}^t} - P_{t-1}^i > 0 \end{cases},$$

where J_{vi}^t is calculated according to expression (3) on the moment of time t. Then on each moment of time t of a period of back-testing $(t = \overline{1,T})$ it is calculated the loss from exchange rates oscillations for the currency briefcase on the whole:

$$\Delta V_t = \sum_{i=1}^{N} \Delta_t^i.$$

2. The comparison of daily meanings VaR_t and corresponding to them actual changes of briefcase cost ΔV_t. The case when the condition is observed $\Delta V_t > VaR_t$ that means that the cost change is negative (loss) and at the same time in absolute meaning exceeds VaR, is called the exceeding of predicted expenses. Then the quantity of cases of exceeding L is calculated.

3. The model adequacy is checked by the parity: $\frac{L}{T} < \alpha$.

8 THE ANALYSIS OF RESULTS OF ESTIMATION OF LOSSES VAR

To estimate VaR it has been used bank currency briefcase that consists of three items in three currencies (USA dollar, euro and Russian ruble). To estimate VaR and test models it has been used the following data: (1) – daily meanings of market rates of dollar USA, euro and Russian ruble from 01.01.2006 – 30.06.2010; (2) – daily data of bank rates from three currencies for a period from 01.06.2006 – 30.06.2010; (3) – daily meanings of open bank currency items in three above mentioned currencies for a period from 01.01.2006 – 30.06.2010.

9 THE ANALYSIS OF THE RESULTS OF ESTIMATION VAR USING DELTA-NORMAL METHOD

On the input of a model the market meanings of exchange rates, bank meanings of exchange rates and bank currency items in briefcase currencies are presented. In order to check

the adequacy of a model the recommendations of Basel committee of bank supervision for different levels of trust (95 %, 99 % and 97 %) have been used. Each three months it is calculated the quantity of mistakes of exchange rate forecast using delta-normal method; it is based on suggestion about normal division of profitability of exchange rates. The depth of retrospective for estimation of standard deviation is 250 days. The depth of retrospective for estimation VaR is also 250 days. The results of verification are brought together in table 1.

According to table 1 we can see that the model for estimation *VaR* of currency briefcase using delta-normal method is inadequate. In order to find the reasons of inadequacy of a model the retrospectives that are used to find estimations *VaR* are checked on normal division according to Pirson criterion. Profitability for USA dollar rate does not have normal division. Profitability for euro is divided close to normal division, and on some periods has normal division.

As a result of retrospective testing of a model while calculating *VaR* for each currency it has been found: if meaning χ^2_{cn} is less according to euro than to USA dollar, then the division of profitability of euro rate is closer to normal than the division of USA dollar. In such case the method gives fewer mistakes of euro forecast than of USA dollar forecast.

It has been found that the losses start exceeding estimation VaR with 95 % level of trust on the period from March 2008 when unpredicted changes of USA dollar and euro rate started to be observed. With 99 % level of trust (which is demanded by Basel committee) the quantity of forecast mistakes starts growing from the end of September 2008. In the period of forced reduction in second – third quarters of 2008 and on the period of 4-th quarter of 2008 – first quarter of 2009 the model for estimation of risk *VaR* using method of historical modeling ceases to be adequate.

10 THE COMPARISON OF METHODS OF ESTIMATION VAR

In table 2 it is represented the results of back-testing for each of models and is calculated the quantity of mistakes according to the results of models work on some periods with duration of 250 days each.

The results of analysis of shortages and advantages of used methods of estimation VaR are gathered in table 3.

11 CONCLUSIONS

Measure *VaR* has shortages and advantages, but it gives possibility to estimate the risk uniquely for each country and each bank. To compare the results of implementation of this methodology on Ukrainian currency market three methods of estimation *VaR* of bank currency briefcase are represented.

The model of risks estimation on the base of delta-normal method has appeared to be inadequate because the assumption about the normal division of currency rates profitability hasn't been made. It is necessary to mark that the division of profitability of euro rate on some periods is close to normal and thus the model of estimation *VaR* of currency item to euro on these periods has appeared to be adequate.

The method of historical modeling has shown satisfactory result only on conditions of stable market situation. It badly adapts to different oscillations on the market and thus today it can't be used on Ukrainian financial market.

The better results of estimation of possible losses have been received using Monte Carlo method. The mistakes in forecasts of possible losses appear only on condition of unpredicted sharp shifts of rate but the model on the base of this method quickly adapts to market changes. To use this method on-line it is necessary to have big calculation capacities that mean vain charges for banks with little market

Table 1. The results of conduction of retrospective testing of calculation VaR using delta-normal method

Period	The results of back - testing					
	95 %		97 %		99 %	
	Quantity of exceeding	% correct forecasts	Quantity of exceeding	% correct forecasts	Quantity of exceeding	% correct forecasts
from 21.03.06 to 21.03.07	13	94,80 %	8	96,80 %	3	98,80 %
from 20.06.06 to 21.06.07	16	93,60 %	7	97,20 %	4	98,40 %
from 22.09.06 to 21.09.07	16	93,60 %	8	96,80 %	6	97,60 %
from 21.12.06 to 21.12.07	21	91,60 %	9	96,40 %	6	97,60 %
from 23.03.07 to 21.03.08	30	88,00 %	20	92,00 %	15	94,00 %
from 24.09.07 to 22.09.08	57	77,20 %	51	79,60 %	40	84,00 %
from 24.12.07 to 22.12.08	65	74,00 %	62	75,20 %	49	80,40 %
from 20.03.08 to 30.03.09	62	75,20%	58	76,80 %	42	83,20 %

Table 2. Comparative analysis of back-testing of a model of risk estimation VaR using different methods

Period		Results of back-testing with 95% level of trust					
		Delta-normal method		Method of historical modeling		Monte Carlo method	
		Quantity of exceeding	% correct forecasts	Quantity of exceeding	% correct forecasts	Quantity of exceeding	% correct forecasts
from 21.03.06	to 21.03.07	13	94,80 %	4	98,40 %	0	100,00 %
from 20.06.06	to 21.06.07	16	93,60 %	5	98,00 %	0	100,00 %
from 22.09.06	to 21.09.07	16	93,60 %	3	98,80 %	0	100,00 %
from 21.12.06	to 21.12.07	21	91,60 %	5	98,00 %	0	100,00 %
from 23.03.07	to 21.03.08	30	88,00 %	16	93,60 %	0	100,00 %
from 22.06.07	to 23.06.08	47	81,20 %	35	86,00 %	3	98,80 %
from 24.09.07	to 22.09.08	57	77,20 %	50	80,00 %	4	98,40 %
from 24.12.07	to 22.12.08	65	74,00 %	74	70,40 %	7	97,20 %
from 20.03.08	to 30.03.09	62	75,20 %	77	69,20 %	7	97,20 %

Table 3. Comparative analysis of work of different methods for risk estimation *VaR*

Method / Criteria	Delta normal	Historical modeling	Method of imitation modeling of Monte Carlo
Estimation	Local	Total	Total
Taking into account historical division	As estimation of normal division	The similar to that in the past	Totally
Taking into account «admissible» volatility	Possible	No	Yes
Assumption about normal division of profitability	Yes	No	No
Estimation of extreme events	Bad	Bad	Possible
Model risk	Can be great	Admissible	High
Volume of retrospective	Average	Very big	Little
Calculation difficulty	Not high	High	Very high
Visualization	Average	High	High
Calculation capacities	Low	Average	High

risk. For these banks it is recommended to use standard approach on the base of fixed coefficients to estimate financial risks. To estimate improbable sharp oscillations of currencies (costs, quoting) it is recommended to use stress-testing that gives representation about the volume of losses in crisis market phenomena.

SPISOK LITERATURY

1. *Лобанов, А. А.* Энциклопедия финансового риск-менеджмента / А. А. Лобанов, А. В. Чугунов. – М. : Альпина Паблишер, 2003. – 786 с.
2. *Jorion, Ph.* Financial risk-management: Second edition / Ph. Jorion. – Hoboken, New Jersey: John Wiley & Sons, 2003. – 708 p.

3. Управління банківськими ризиками : навч. посібник / [Примостка Л. О., Чуб П. М., Карчева Т. Г. та ін.]; за ред. Примостки О. Л. – К. : КНЕУ, 2007. – 600 с.
4. Методичні вказівки з інспектування банків «Система оцінки ризиків» / Затв. постановою правління Національного банку України від 15.03.2004 № 104.– 2004. – 43 с.
5. Базельский комитет по банковскому надзору. Международная конвергенция измерения капитала и стандартов капитала: новые подходы. – Базель, 2004. – http://www.cbr.ru
6. *Яблоков, А. І.* Методика оцінювання та управління валютним ризиком VaR / А. І. Яблоков // Економіко-математичне моделювання соціально-економічних систем. – 2007. – № 13. – С. 121–128.
7. *Милосердов, А. А.* Рыночные риски: формализация, моделирование, оценка качества моделей / А. А. Милосердов, Е. Б. Герасимова. – Тамбов : Изд-во Тамбовского гос. техн. ун-та, 2004. – 116 с.

Стаття надійшла до редакції 28.10.2013.

Трофимчук А. Н.[1], Кожуховская О. А.[2], Бидюк П. И.[3], Кожуховский А. Д.[4]
[1]Д-р техн. наук, професор, Институт телекоммуникаций и информационных технологий НАН Украины, Украина
[2]Канд. техн.наук, ст. преп., Черкасский государственный технологический университет, Украина
[3]Д-р техн. наук, профессор, Институт прикладного системного анализа Национального технического университета Украины «КПИ»,Украина
[4]Д-р техн. наук, профессор Черкасский государственный технологический университет, Украина

ОЦЕНКА РЫНОЧНОГО РИСКА В УКРАИНЕ ПО МЕТОДОЛОГИИ VAR

В работе рассматривается возможность применения методов оценивания меры риска VaR к банковскому валютному портфелю с использованием таких методов: дельта-нормальный, исторический и имитационного моделирования. Приемлемые по качеству результаты прогнозирования возможных потерь получены по методу Монте-Карло, который гипотетически может учитывать возможные изменения курсов валют на рынке. Установлено, что ошибки прогнозов возможных потерь могут возникать только вследствие резких непредвиденных изменений курсов валют.

Ключевые слова: мера риска VaR, банковский валютный портфель, историческое моделирование, имитационное моделирование по методу Монте-Карло, дельта-нормальный метод, украинский валютный рынок.

Трофимчук О. М.[1], Кожухівська О. А.[2], Бідюк П. І.[3], Кожухівський А. Д.[4]
[1]Д-р техн. наук, професор, Інститут телекомунікацій та інформаційних технологій НАН Україна
[2]Канд. техн.наук, ст.викл., Черкаський державний технологічний університет, Україна
[3]Д-р техн. наук, професор, Інститут прикладного системного аналізу Національного технічного університету України «КПІ»,Україна
[4]Д-р техн. наук, професор, Черкаський державний технологічний університет, Україна

ОЦІНКА РИНКОВОГО РИЗИКУ В УКРАЇНІ ЗА МЕТОДОЛОГІЄЮ VAR

В роботі розглядається можливість застосування методів оцінювання міри ризику VaR для банківського валютного портфеля з використанням таких методів: дельта-нормальний, історичний та імітаційного моделювання. Прийнятні за якістю результати прогнозування втрат отримано за методом Монте-Карло, який гіпотетично може враховувати можливі зміни курсів валют на ринку. Встановлено, що похибки прогнозів можливих втрат виникають лише за наявності непередбачуваних різких змін курсу валют.

Ключові слова: міра ризику VaR, банківський валютний портфель, історичне моделювання, імітаційне моделювання за методом Монте-Карло, дельта-нормальний метод, український валютний ринок.

REFERENCES

1. Lobanov A. A., Chugunov A. V. E'nciklopediya finansovogo risk-menedzhmenta. Moscow, Alpina Pablisher, 2003, 786 p.
2. Jorion Ph. Financial risk-management: Second edition. Hoboken, New Jersey, John Wiley & Sons, 2003, 708 p.
3. Primostka L. O., Chub P. M., Karcheva T.G. ta in.; za red. Primostky L.O Upravlinnya bankivskymy rysykamy : navch. Posibnyk. Kyiv, KNEU, 2007, 600 p.
4. Metodychni vkazivky z inspektuvannya bankiv «Systema ozinky rysykiv». Zatv. postanovoiu pravlinnia Nazionalnogo banku Ukrainy vid 15.03.2004 № 104, 2004, 43 p.
5. Bazelskij komitet po bankovskomu nadzoru. Mezhdunarodnaya konvergensiya izmereniya kapitala i standartov kapitala: novy'e podhody', Bazel, 2004. http://www.cbr.ru
6. Yablokov A. I. Metodyka ozinyuvannya ta upravlinnya valiutnim ryzykom VaR, E'konomiko-matematychne modeliuvannia sotsialno-ekonomichnyh system, 2007, No. 13, pp. 121–128.
7. Miloserdov A. A., Gerasimova E. B. Ry'nochny'e riski: formalizaciya, modelirovanie, ocenka kachestva modelej. Tambov: Izd-vo Tambovskogo gos. tehn. un-ta, 2004, 116 p.

УДК 004.02

Чопоров С. В.

Канд. техн. наук, доцент, Запорожский национальный университет, Украина, E-mail: choporov@list.ru

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ В МЕТОДЕ КОНЕЧНЫХ ЭЛЕМЕНТОВ

В работе рассмотрена проблема использования параллельных вычислений в методе конечных элементов. Рассмотрены особенности построения матрицы жесткости и решения полученной системы линейных уравнений на базе параллельных вычислений. Проведены вычислительные эксперименты, результаты которых использованы для анализа эффективности предложенных подходов.

Ключевые слова: метод конечных элементов, параллельные вычисления, математическая модель, вычислительный метод.

ВВЕДЕНИЕ

Проектирование современных деталей и конструкций связано с необходимостью применения вычислительных методов для решения задач высокой сложности. В разработке САПР для машиностроения наиболее активно применяется один из наиболее распространенных вычислительных методов – метод конечных элементов, который находит свои приложения в широком спектре задач: от анализа напряженно деформированного состояния деталей, конструкций и строений, переноса тепла, гидродинамики до анализа микроэлектромеханических систем [1–3].

Сложность многих деталей и конструкций приводит к возникновению на подготовительных этапах сложных геометрических моделей – математических моделей соответствующих геометрических областей. Последующее применение методов дискретизации [4] с требованием учета особенностей геометрической модели и физической постановки задачи (например, сгущение сетки конечных элементов в областях контакта, трещин, учет многослойности и другие) может приводить к дискретным моделям с большим числом конечных элементов. В результате для моделирования необходимы значительные вычислительные ресурсы и, как следствие, временные затраты.

Естественным решением задачи минимизации вычислительных затрат является оптимизация числа элементов в дискретных моделях. Однако, в тех случаях, когда уменьшение числа конечных элементов приводит к неоправданной потере точности модели, одним из путей оптимизации временных затрат является использование параллельных вычислений. Стремительное увеличение числа функциональных устройств (процессоров, ядер) в современных вычислительных системах актуализирует данное направление исследований.

Таким образом, целью работы является разработка параллельных алгоритмов для основных этапов метода конечных элементов и анализ результатов их внедрения путем проведения вычислительных экспериментов.

1 ТИПЫ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ И ОСНОВНЫЕ ТЕХНОЛОГИИ ДОСТИЖЕНИЯ ПАРАЛЛЕЛЬНОСТИ

Одной из наиболее распространенных классификаций архитектуры вычислительных систем является систематика Флинна [5], которая основана на анализе способов взаимодействия функциональных вычислительных устройств и обрабатываемых данных. В результате различаются основные четыре типа вычислительных систем [5–7]:

– SISD (Single Instruction, Single Data) – вычислительные системы, в которых используется единственное функциональное устройство для исполнения одного потока инструкций с одиночным потоком данных;

– SIMD (Single Instruction, Multiple Data) – системы с одним потоком инструкций, в которых одна инструкция может применяться к векторам данных (нескольким потокам данных);

– MISD (Multiple Instruction, Single Data) – системы с множественным потоком инструкций и одним потоком данных;

– MIMD (Multiple Instruction, Multiple Data) – вычислительные системы с множественным потоком инструкций и множественным потоком данных.

Большинство современных последовательных вычислительных систем можно отнести к первым двум типам. Такие компьютеры оснащены одним функциональным устройством и поддерживают различные наборы SIMD-инструкций (например, MMX, SSE, SSE2 и другие). Параллельность (квазипараллельность) в таких системах является реализацией многозадачности в планировщике задач операционной системы.

К типу MISD можно отнести отказоустойчивые вычислительные системы, которые выполняют инструкции с функциональной избыточностью с целью обнаружения ошибок. При этом ряд авторов [5] называют данный тип теоретическим, отмечая отсутствие устройств такого типа.

К типу MIMD относятся системы с множеством функциональных устройств: мультипроцессоры (многопроцессорные и многоядерные системы) и мультикомпьютеры (распределенные вычислительные системы, кластеры).

На современном рынке вычислительных систем в наиболее доступном ценовом диапазоне широко представлены мультипроцессорные системы с общей памятью, которые содержат от 2 до 16 функциональных устройств. Такие вычислительные системы можно использовать для начальной отладки параллельных алгоритмов и оценки их эффективности. Для программирования в таких системах наиболее распространенными являются решения на базе потоков (например, POSIX Threads, Boost.Thread и другие) или стандарта OpenMP (Open Multi-Processing) [8, 9].

2 ПАРАЛЛЕЛЬНАЯ РЕАЛИЗАЦИЯ МЕТОДА КОНЕЧНЫХ ЭЛЕМЕНТОВ

Основная идея метода конечных элементов состоит в том, что любую искомую непрерывную величину (например, температуру, перемещения и другие) можно аппроксимировать при помощи кусочно-непрерывных функций, определенных на элементарных геометрических областях – конечных элементах, образующих дискретную модель области. Для определения значений искомой непрерывной величины в узлах дискретной модели строится система линейных алгебраических уравнений, матрица которой носит название глобальной матрицы жесткости.

Процесс построения глобальной матрицы жесткости называется ансамблированием и является суммой, подчиненной особым правилам, локальных матриц жесткости – результатов интегрирования дифференциального уравнения в матричной форме. Размер глобальной матрицы жесткости равен произведению числа узлов в дискретной модели на число степеней свободы в узле (определяется типом задачи).

Таким образом, в случае достаточно сложных дискретных моделей, этапы построения глобальной матрицы жесткости и последующего решения СЛАУ являются наиболее требовательными к вычислительным ресурсам. Следовательно, внедрение параллельных вычислений при выполнении этих этапов способно дать наиболее значимый результат с точки зрения экономии времени на проектирование и моделирование.

Первым шагом после этапов загрузки и инициализации исходных данных параллельной реализации метода конечных элементов является определение величины K – числа доступных функциональных устройств (рис. 1). Данная величина является основой для определения количества параллельных потоков. В методе конечных элементов при вычислении глобальной матрицы жесткости в случае мультипроцессорных систем операции ввода-вывода используются только, если недостаточно основной памяти для хранения исходных данных и самой глобальной матрицы жесткости. То есть, при использовании сжатого способа хранения глобальной матрицы жесткости (которая, как правило, является сильно разреженной), время ожидания операций ввода-вывода, в зависимости от сложности задачи, будет минимальным, следовательно, на практике величина K также определяет количество параллельных вычислительных потоков.

Так как ансамблирование является обобщением суммы, то каждый из K параллельных потоков должен производить ансамблирование N / K элементов (N – количество элементов в дискретной модели). При этом все данные (включая переменные циклов), связанные с интегрированием на конечном элементе должны быть локальными данными параллельных потоков (быть доступными только для потока-обладателя; схема приведена на рис. 1).

Шаг обновления глобальной матрицы жесткости, учитывая сжатый формат хранения данных, является критическим с точки зрения потокобезопасности. То есть, доступ остальным потокам к глобальной матрице жесткости должен блокироваться пока один поток модифицирует ее элементы.

Время, необходимое для построения глобальной матрицы жесткости при помощи параллельного вычислительного процесса, можно оценить величиной T_p

$$T_p = N\left(I / K + m\right) + K \cdot t_{thread},$$

где I – процессорное время, необходимое для интегрирования исходного дифференциального уравнения на одном конечном элементе, m – процессорное время, необходимое для ансамблирования результатов интегрирования на одном элементе, t_{thread} – время, необходимое для инициализации одного вычислительного потока.

Аналогично, время, необходимое для построения глобальной матрицы жесткости на базе последовательного алгоритма, можно оценить величиной T_s

$$T_s = N\left(I + m\right).$$

Применение параллельных вычислений имеет смысл, если $T_s > T_p$:

$$N\left(I + m\right) > N\left(I / K + m\right) + K \cdot t_{thread},$$

в результате упрощения которой получим следующую оценку

$$N \cdot I > \frac{K^2 \cdot t_{thread}}{K - 1}.$$

Таким образом, например, для двух параллельных потоков, произведение количества элементов на время интегрирования одного элемента должно быть вчетверо большим времени инициализации параллельных потоков, что может не выполняться при малом числе конечных элементов.

Для решения полученных в результате построения глобальной матрицы жесткости СЛАУ на практике, как правило, используются итерационные методы (напри-

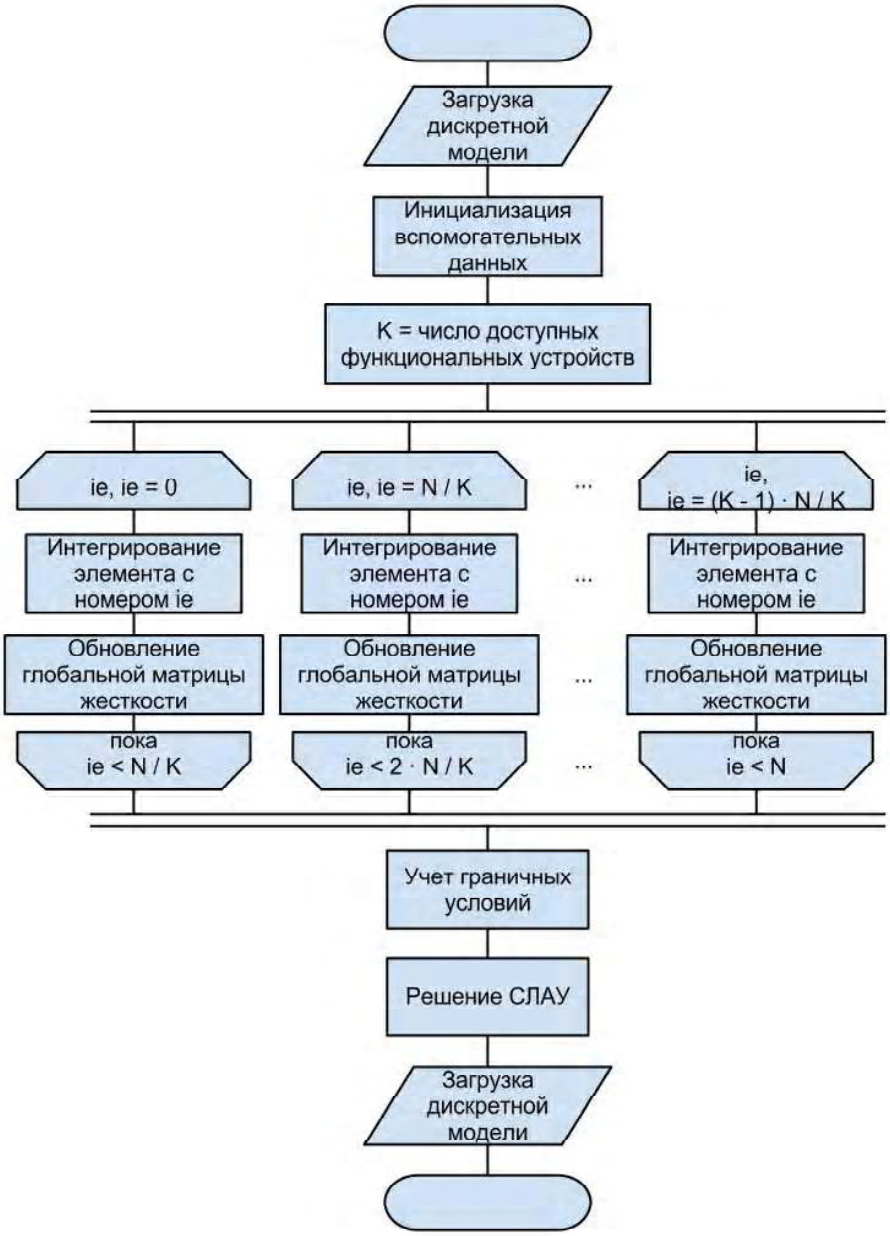


Рис. 1. Схема метода конечных элементов на базе параллельных вычислений

мер, метод сопряженных градиентов [10, 11] и другие). Их ключевой особенностью является то, что результаты текущей итерации являются исходными данными для следующей. Следовательно, для сокращения времени поиска решения необходимо ускорять внутренние операции метода: произведение матрицы на вектор, поэлементное произведение векторов, вычисления невязки и другие.

С учетом того, что в результате ансамблирования глобальная матрица жесткости получается сильно разреженной, следовательно, схема вычисления произведения матрицы на вектор в K параллельных потоков примет вид, приведенный на рис. 2. Общими для потоков данными являются матрица A, размер которой $D \times D$, вектор x, вектор-результат r и число K. Так как вычислительными потоками обновляются разные элементы вектора r, то отсутствует необходимость в блокировке для синхронизации потоков. Счетчики циклов i и j, переменная t – локальные для каждого вычислительного потока. Во внутренних циклах по j вычисления производятся только для ненулевых элементов i-й строки матрицы A.

3 ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

В качестве тестового стенда использовалась программно-аппаратная конфигурация с процессором Intel Core i3-380 (2,53 ГГц, 2 физических ядра + 2 виртуальных ядра) с 3 ГБ ОЗУ под управлением операционной системы openSUSE 12.2 (компилятор gcc 4.7), для создания потоков и управления ими – библиотеки стандарта openMP.

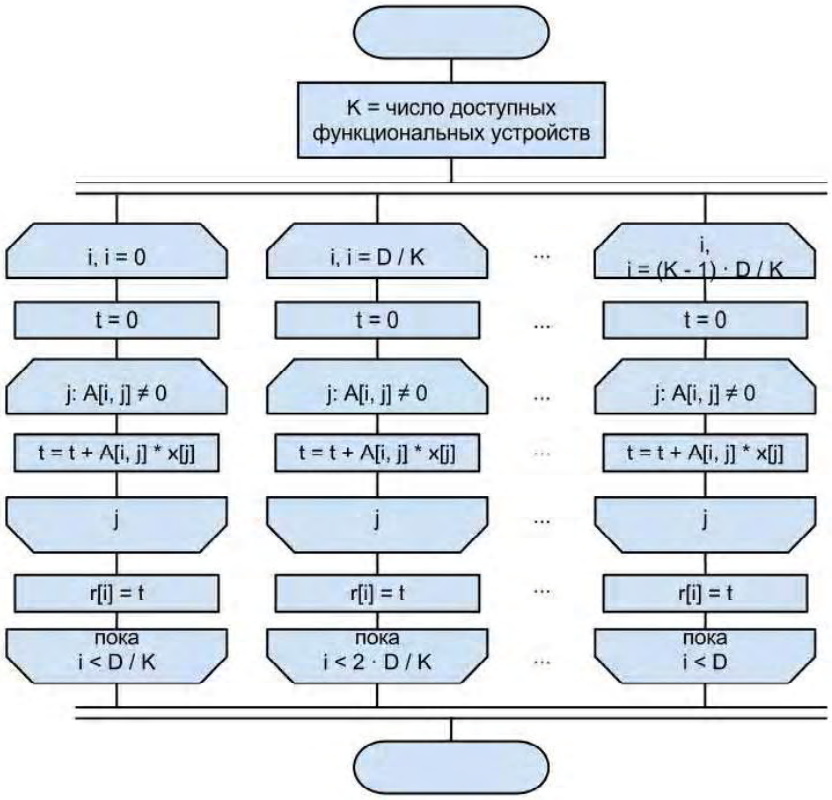


Рис. 2. Схема параллельного произведения разреженной матрицы на вектор

1. Задача Дирихле

Рассмотрим решение уравнения Лапласа с условиями Дирихле.

$$\frac{\partial^2 U}{\partial x^2}+\frac{\partial^2 U}{\partial y^2}+\frac{\partial^2 U}{\partial z^2}=0,$$
$$U_{\delta\Omega}=f(x,y,z).$$

В качестве области Ω рассмотрим косозубое зубчатое колесо (дискретная модель: 49500 узлов, 40920 шестигранных конечных элементов, рис. 3). Функция $f(x, y, z)$ пусть имеет вид

$$f(x,y,z)=\begin{cases}0, \text{в узлах центрального отверстия},\\ 4\sin^2(5\varphi), \text{на поверхности зубьев},\end{cases}$$

где φ – угол поворота полярного радиуса до соответствующего узла в плоскости Oxy.

Формула для вычисления локальной матрицы жесткости $[K^e]$ примет вид

$$\left[K^e\right]=\iiint\left(\begin{Bmatrix}\frac{\partial H_1}{\partial x}\\ \frac{\partial H_2}{\partial x}\\ \dots\\ \frac{\partial H_8}{\partial x}\end{Bmatrix}\begin{bmatrix}\frac{\partial H_1}{\partial x} & \frac{\partial H_2}{\partial x} & \dots & \frac{\partial H_8}{\partial x}\end{bmatrix}+\right.$$
$$+\begin{Bmatrix}\frac{\partial H_1}{\partial y}\\ \frac{\partial H_2}{\partial y}\\ \dots\\ \frac{\partial H_8}{\partial y}\end{Bmatrix}\begin{bmatrix}\frac{\partial H_1}{\partial y} & \frac{\partial H_2}{\partial y} & \dots & \frac{\partial H_8}{\partial y}\end{bmatrix}+$$
$$\left.+\begin{Bmatrix}\frac{\partial H_1}{\partial z}\\ \frac{\partial H_2}{\partial z}\\ \dots\\ \frac{\partial H_8}{\partial z}\end{Bmatrix}\begin{bmatrix}\frac{\partial H_1}{\partial z} & \frac{\partial H_2}{\partial z} & \dots & \frac{\partial H_8}{\partial z}\end{bmatrix}\right)dxdydz,$$

где H_i (i = 1, 8) – функции формы конечного элемента.

В результате вычислительного эксперимента получены результаты, приведенные в табл. 1 (время в секундах).

2. Исследование прогиба трехслойной круглой оболочки под равномерной нагрузкой

Дана защемленная по контуру трехслойная оболочка, радиус которой 0,4 м, толщина – 0,01 м. Внутренний слой: толщина – 0,008 м, модуль Юнга – E_i = 72017,3327 МПа, коэффициент Пуассона – ν_i = 0,2999518536. Внешние

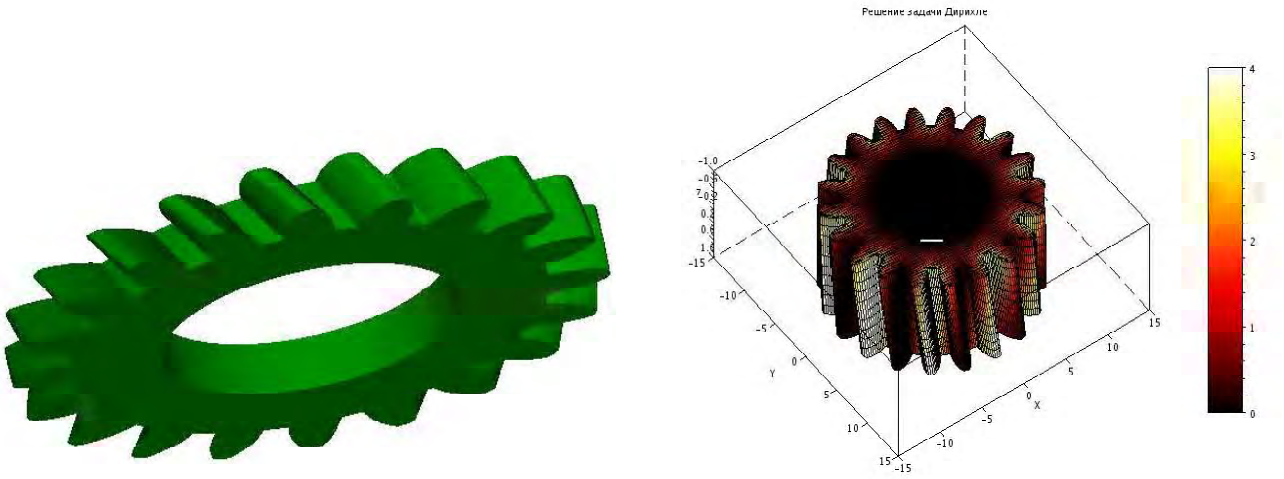


Рис. 3. Задача Дирихле: косозубое зубчатое колесо

слои: толщина – 0,001 м, модуль Юнга – E_o = 203200 МПа, коэффициент Пуассона – ν_o = 0,27. На оболочку действует равномерная нагрузка q = –0,05 МПа.

Для решения построена дискретная модель сектора оболочки (рис. 4) на базе десяти слоев шестигранных конечных элементов (59213 узлов и 52240 элементов), толщина которых 0,001 м, что позволяет рассматривать данную задачу в трехмерной постановке, используя соответствующие слоям упругие константы.

Формула для вычисления локальной матрицы жесткости $[K^e]$ примет вид

$$\left\lfloor K^e \right\rfloor = \iiint B^T DB dx dy dz,$$

где

$$B = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 \\ 0 & \frac{\partial}{\partial y} & 0 \\ 0 & 0 & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial x} \end{bmatrix} \times$$

$$\times \begin{bmatrix} H_1 & H_2 & \dots & H_8 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & H_1 & H_2 & \dots & H_8 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & H_1 & H_2 & \dots & H_8 \end{bmatrix},$$

Таблица 1. Решение задачи Дирихле

Этап	Количество вычислительных потоков		
	1	2	4
	время, с		
Построение глобальной матрицы жесткости	3,07639	1,56709	1,25949
Решение СЛАУ, 73 итерации	0,304023	0,240893	0,264904

$$D = \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} \times$$

$$\times \begin{bmatrix} 1 & \frac{\nu}{1-\nu} & \frac{\nu}{1-\nu} & 0 & 0 & 0 \\ \frac{\nu}{1-\nu} & 1 & \frac{\nu}{1-\nu} & 0 & 0 & 0 \\ \frac{\nu}{1-\nu} & \frac{\nu}{1-\nu} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} \end{bmatrix}$$

$$E = \begin{cases} E_o, \; z \geq 0{,}009 \text{ или } z \leq 0{,}001, \\ E_i, \; 0{,}001 \leq z \leq 0{,}009, \end{cases}$$

$$\nu = \begin{cases} \nu_o, \; z \geq 0{,}009 \text{ или } z \leq 0{,}001, \\ \nu_i, \; 0{,}001 \leq z \leq 0{,}009. \end{cases}$$

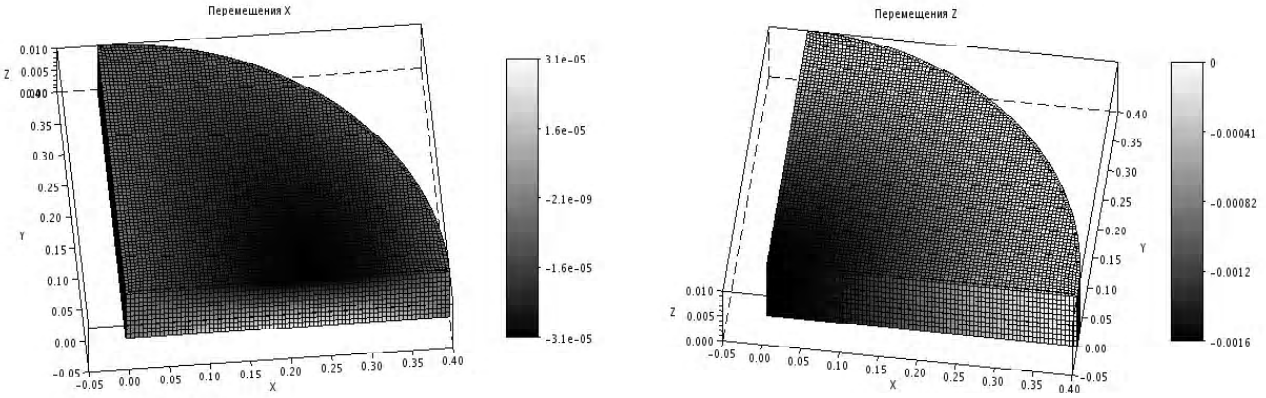


Рис. 4. Прогиб трехслойной оболочки

В результате вычислительного эксперимента получены результаты, приведенные в табл. 2 (время в секундах).

Таблица 2. Исследование прогиба трехслойной круглой оболочки

Этап	Количество вычислительных потоков		
	1	2	4
	время, с		
Построение глобальной матрицы жесткости	83,7293	41,5063	31,7687
Решение СЛАУ, 6952 итерации	298,287	191,299	189,134

ВЫВОДЫ

Таким образом, вычислительные эксперименты показали, что использование параллельных вычислений в методе конечных элементов позволяет сократить время на проектирование и моделирование деталей и конструкций. Из таблиц видно незначительное преимущество во времени при использовании четырех вычислительных потоков относительно двух, что может быть связано с использованием двух виртуальных ядер (которые не обладают аналогичной двум физическим ядрам производительностью). Реализация разработанных алгоритмов на базе библиотек стандарта openMP позволяет рассматривать полученную в ходе исследования кодовую базу в качестве основы для построения кластерной реализации. Перспективным направлением исследований в данном направлении является изучение влияния количества потоков на производительность (или затраченное время). Также важным направлением является изучение производительности многопоточных вычислений в зависимости от количества узлов/элементов в сетке, особенно в системах с раздельной памятью (вычислительных кластерах), в которых существуют значительные затраты на синхронизацию вычислительных потоков.

СПИСОК ЛИТЕРАТУРЫ

1. *Zienkiewicz, O. Z.* The Finite Element Method. V. 1: The Basis / O. Z. Zienkiewicz, R. L. Taylor. – Oxford : Butterworth-Heinemann, 2000. – 689 p.
2. *Smith, I. M.* Programming the finite element method / I. M. Smith, D. V. Griffits. – Chichester : Wiley, 2004. – 628 p.
3. *Ahmad, A.* Parallel Programming in the Finite Element Methos / Atique Ahmad // In. Proceedings of Failure of engineering Materials and Structures, UET Taxila, October 2007. – P. 87–93.
4. *Thompson, J. F.* Hand book of grid generation / Joe F. Thompson, Bharat K. Soni, Nigel P. Weatheril. – New York : CRC Press, 1999. – 1200 p.
5. *Гергель, В. П.* Основы параллельных вычислений для многопроцессорных вычислительных систем / В. П. Гергель, Р. Г. Стронгин. – Нижний Новгород : ННГУ им. Н.И. Лобачевского, 2003. – 184 с.
6. *Корнеев, В. В.* Параллельные вычислительные системы / В. В. Корнеев. – М. : Нолидж, 1999. – 320 с.
7. *Hwang, K.* Scalable Parallel Computing: Technology, Architecture, Programming / Kai Hwang, Zhiwei Xu. – New York : McGraw-Hill, 1998. – 802 p.
8. *Антонов, А. С.* Параллельное программирование с использованием технологии OpenMP : учебное пособие / А. С. Антонов. – М. : МГУ, 2009. – 77 с.
9. *Chapman, B.* Using OpenMP : portable shared memory parallel programming / Barbara Chapman. – Cambridge : MIT, 2008. – 353 p.
10. *Notay, Y.* Flexible Conjugate Gradients / Yvan Notay // SIAM Journal on Scientific Computing. – 2001. – Volume 22, Issue 4. – P. 1444–1460.
11. *Knyazev, A. V.* Steepest Descent and Conjugate Gradient Methods with Variable Preconditioning / Andrew V. Knyazev, Ilya Lashuk // SIAM Journal on Matrix Analysis and Applications. – 2008. – Volume 29, Issue 4. – P. 1267–1280.

Стаття надійшла до редакції 15.02.2013.

Чопоров С. В.
Канд. техн. наук, доцент, Запорізький національний університет, Україна

ВИКОРИСТАННЯ ТЕХНОЛОГІЙ ПАРАЛЕЛЬНИХ ОБЧИСЛЕНЬ У МЕТОДІ СКІНЧЕННИХ ЕЛЕМЕНТІВ

В роботі розглянута проблема впровадження паралельних обчислень у метод скінченних елементів. Розглянуто особливості побудови матриці жорсткості та вирішення отриманої системи лінійних рівнянь на базі паралельних обчислень. Проведено обчислювальні експерименти, які використані для аналізу ефективності запропонованих підходів.

Ключові слова: метод скінченних елементів, паралельні обчислення, математична модель, обчислювальний метод.

Choporov S. V.
Ph.D. Candidate, associate Professor, Zaporizhzhya National University, Ukraine

PARALLEL COMPUTING TECHNOLOGIES IN THE FINITE ELEMENT METHOD

Nowadays engineers and researchers are faced with solving very complex problems in a mathematical modeling and design. Numerical analysis naturally finds applications in all fields of engineering and the physical sciences. The finite element method is a powerful tool for the numerical simulation of a wide range of problems. Implementation of the finite element method in CAD systems on the basis of modern computers allows researchers to solve large scale problems.

The finite element method uses a discretization of a continuum domain into a mesh as the starting point. The discretization of complex domains may give large numbers of elements, thus increasing the requirements of computer memory and speed.

Modern parallel computers use multiple processing elements simultaneously to solve a problem. Thus implementation of parallel computing into the finite element method is urgent direction of the research.

In the finite element method for the numerical solution of partial differential equations, the stiffness matrix represents the system of linear equations that must be solved in order to ascertain an approximate solution to the differential equation. Therefore the article describes parallel algorithms for assembly of stiffness matrix and for solution of linear equations.

Also this article contains two numerical experiments: 1) Dirichlet problem on the complex domain (gear); 2) Elasticity problem of three-layer shell.

In the end of the article author compares the speed of a solution of these problems with one, two and four parallel cores and makes discussion about the effectiveness of parallel finite element processing.

Keywords: finite element method, parallel computing technologies, mathematical model, numerical method.

REFERENCES

1. Zienkiewicz O. Z., Taylor R.L. The Finite Element Method. V. 1: The Basis. Oxford, Butterworth-Heinemann, 2000, 689 p.
2. Smith I. M., Griffits D. V. Programming the finite element method. Chichester, Wiley, 2004, 628 p.
3. Ahmad A. Parallel Programming in the Finite Element Methos, *In. Proceedings of Failure of engineering Materials and Structures*, UET Taxila, October 2007, pp. 87–93.
4. Thompson J. F., Bharat K. Soni, Nigel P. Weatheril. Hand book of grid generation. New York, CRC Press, 1999, 1200 p.
5. Gergel' V. P., Strongin R. G. Osnovy parallel'nyx vychislenij dlya mnogoprocessornyx vychislitel'nyx system. Nizhnij Novgorod, NNGU im. N. I. Lobachevskogo, 2003, 184 p.
6. Korneev V. V. Parallel'nye vychislitel'nye sistemy. Moscow, Nolidzh, 1999, 320 p.
7. Hwang K., Zhiwei Xu. Scalable Parallel Computing: Technology, Architecture, Programming. New York, McGraw-Hill, 1998, 802 p.
8. Antonov A. S. Parallel'noe programmirovanie s ispol'zovaniem texnologii OpenMP: Uchebnoe posobie. Moscow, MGU, 2009, 77 p.
9. Chapman B. Using OpenMP : portable shared memory parallel programming. Cambridge, MIT, 2008, 353 p.
10. Notay Y. Flexible Conjugate Gradients, *SIAM Journal on Scientific Computing*, 2001, Volume 22, Issue 4, pp. 1444–1460.
11. Knyazev A.V., Lashuk Ilya Steepest Descent and Conjugate Gradient Methods with Variable Preconditioning, *SIAM Journal on Matrix Analysis and Applications*, 2008, Volume 29, Issue 4, pp. 1267–1280.

УДК 681.5.015

Шевцов Д. В.

Канд. техн. наук, доцент, Донецкий национальный университет, Украина, E-mail: shedmival@mail.ru

ОПРЕДЕЛЕНИЕ ОРИЕНТАЦИИ ЭЛЕМЕНТАРНЫХ СОСТАВЛЯЮЩИХ МОДЕЛЕЙ ЗНАКОВ, ПОДЛЕЖАЩИХ АВТОМАТИЧЕСКОМУ ИМЕНОВАНИЮ НА МНОЖЕСТВЕ АТОМАРНЫХ ЭЛЕМЕНТОВ

Автоматическое моделирование, именование и опознавание знаков цифровых бинарных изображений произвольной природы является актуальной научно-практической задачей, находящей свое приложение во многих областях внедрения информационных технологий, в частности, при обработке и анализе электронных документов. Статья посвящена конструктивному определению структурных составляющих моделируемых изображений, прошедших этап утоньшения при предварительной обработке, и их свойств, которые позволят осуществлять автоматическое выявление указанных объектов с целью последующего формирования описания именуемых и распознаваемых знаков на дискретном множестве атомарных элементов.

Ключевые слова: автоматическое моделирование, цифровые бинарные изображения, распознавание, кратчайший путь.

ВВЕДЕНИЕ

Согласно [1], автоматический анализ, именование и распознавание цифровых бинарных изображений произвольной природы, осуществляемые в терминах дискретного множества атомарных элементов (АЭ), предполагает, в частности, решение такой актуальной научно-практической задачи, как определение фрагментов GT-объектов (от англ. «generative trajectory» – «порождающие траектории») бинарных растровых цифровых изображений (РЦИ) знаков, сопоставимых со структурными элементами модели, формируемой для дальнейшего распознавания исходных объектов [2].

Данный подход аналогичен решению задачи выбора непроизводных элементов при сегментации изображений.

Однако отличается от указанного тем, что не предполагает априорного анализа предметной области решаемой задачи распознавания и множества репрезентативных представителей классов распознаваемых объектов. В работе [3] обоснована целесообразность решения задачи сегментации в терминах свойств дискретных множеств путем приведения ее к задаче автоматической декомпозиции знаков на дискретном множестве АЭ. При этом в качестве примитивов используются конструктивно определенные кратчайшие пути (КП) [3], что предполагает возможность их автоматического выявления на GT-объектах [4], результирующих утоньшение исходных знаков на этапе предварительной обработки.

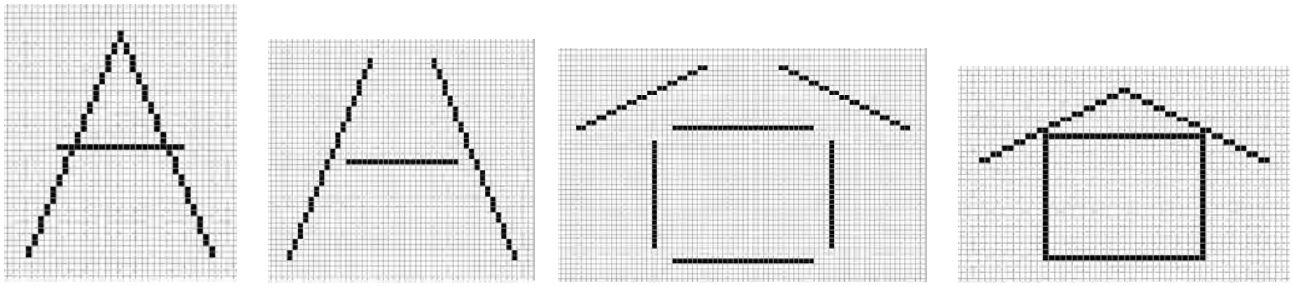
При этом в работе [3] отмечено, что автоматическая декомпозиция знаков с целью их последующего моделирования и именования предполагает необходимость определения различных свойств КП, учитывающих их качественные характеристики и отражающие такие свойства фрагментов знаков, как выпуклость, вогнутость и прочие [5, 6]. Настоящая работа посвящена конструктивному определению указанных характеристик фрагментов GT-объектов знаков, используемых в дальнейшем при формировании моделей обрабатываемых системой распознавания исходных изображений. Использование предложенных понятий позволит осуществлять автоматическую декомпозицию и моделирование знаков произвольной природы на дискретном множестве АЭ с учетом свойств выявляемых КП, существенных для их различения, анализа и именования [1].

1 ОСНОВНЫЕ ТЕОРЕТИЧЕСКИЕ ПОЛОЖЕНИЯ ДЕКОМПОЗИЦИИ GT-ОБЪЕКТОВ ЗНАКОВ НА МНОЖЕСТВЕ АТОМАРНЫХ ЭЛЕМЕНТОВ

При проведении последующих рассуждений будем полагать, согласно [1–4], что автоматическому моделированию и именованию подлежат образы GT-объектов бинарных растровых цифровых изображений знаков произвольной природы [2]. В соответствии с [3], актуальной является проблема конструктивного определения в терминах свойств дискретного множества АЭ объектов, допускающих однозначную декомпозицию GT-объектов, аналогичная описанному в [1, 5] заданию множества непроизводных элементов в рамках структурных методов распознавания.

На рис. 1 приведены примеры исходных бинарных РЦИ и множеств непроизводных элементов, априорно заданных для каждого класса распознаваемых объектов.

В рамках сделанного в [1] предположения о том, что словарь W имен моделируемых объектов и множество B изображений знаков обучающей выборки могут быть изначально пусты, указанное задание множеств непроизводных элементов затруднительно. С целью конструктивного определения в терминах свойств множеств АЭ минимальных составляющих автоматически формиру-



а) *б)* *в)* *г)*

Рис. 1. Исходные растровые цифровые изображения (а, в) и традиционно используемые множества непроизводных элементов для их моделирования (б, г)

емых математических моделей (ММ) образов GT-объектов именуемых знаков изображений, последующие рассуждения проведем в рамках гипотезы о способе генерации соответствующих изображений, изложенной в [2].

Будем полагать, что изображение произвольного знака открытого алфавита представляет собой визуализированный на плоскости регистрации след некоторого устройства фиксации [2], осуществляющего движения по траекториям, каждая из которых взаимно однозначно соответствует непрерывной кривой без самопересечений, заданной на подпространстве пространства E_2, множество $M_j^i = \{\rho_k^{ij}(x, y) = 0\}_{k=1}^{N_{ij}}$, $i, j \in N$, которых является моделью генерируемого изображения [2].

В системах распознавания, обзор методологий проектирования которых представлен в [1], осуществляется выявление образов указанных кривых и их фрагментов, трактуемых либо как признаки [5], либо как непроизводные элементы [6]. При этом первичным является эвристический анализ изображений обучающей выборки [1]. Следует заметить, что априорное задание множества образов кривых моделей сгенерированных изображений в общем случае проблематично [1, 5]. Также затруднительно, используя категории и объекты всюду плотных множеств [5, 6], осуществлять автоматическое выявление на анализируемых образах изображений фрагментов, однозначно либо с некоторой точностью сопоставимых с выбранными элементарными составляющими моделей [1], что негативно влияет на точность распознавания [5].

В соответствии с этим, в работах [3, 4] предложено множество образов кривых, восстановленных по образу изображения, каждый из которых соответствует единственной траектории движения регистрирующей части [2] устройства фиксации следа, осуществленного в процессе генерации, рассматривать в качестве модели изображения знака. Указанное множество образов кривых определено в [2] как GT-объект – модель образа изображения знака.

Согласно проведенным рассуждениям, одной из задач исследования данной работы является определение на дискретном множестве АЭ образов кривых GT-объекта, моделирующих траектории, порождающие исходное изображение знака [1, 2]. Для ее решения рассмотрим множество АЭ $A = \{\alpha_h\}_{h=1}^{H}$, где $\alpha_h = \alpha(i_h, j_h)$, $i_h \in \{1, 2, ..., I\}$, $j_h \in \{1, 2, ..., J\}$, $H = I \cdot J$ [ ], на котором $\forall \alpha_a, \alpha_b \in A$ введена метрика:

$$\rho(\alpha_a, \alpha_b) = |i_a - i_b| + |j_a - j_b|. \quad (1)$$

и определены четыре типа связок s_m, $m \in \hat{M} = \{1, 2, 3, 4\}$ [3].

Как показано в [2, 4], на множестве АЭ образом непрерывной кривой без самопересечений, заданной в терминах свойств всюду плотных множеств, является путь [3] как конечное вполне упорядоченное множество связных связок – $L(\alpha_a, \alpha_b) = \{(\alpha_h, \alpha_{h+1})_{m_h}\}_{h=1}^{n}$, $m_h \in \hat{M}$, где $\alpha_1 = \alpha_a$, $\alpha_{n+1} = \alpha_b$ – соответственно начальный и конечный АЭ пути, $(\alpha_h, \alpha_{h+1})_{m_h}$ – связка типа $m_h \in \hat{M}$, $h = \overline{1, n}$. При этом, согласно [3], АЭ α_r, $\forall r \in \{2, 3, ..., n\}$, имеет ровно два связных АЭ из множества $\Lambda(L(\alpha_a, \alpha_b)) = \{\alpha_k\}_{k=1}^{n+1}$, а АЭ α_1 и α_{n+1} – не более двух.

Таким образом, пути на множестве АЭ моделируют образы GT-объектов, в частности, результирующих утоньшение [2] или скелетизацию [5] образов РЦИ знаков различных алфавитов, что соответствует изложенной гипотезе о способе генерации изображений знаков и подтверждается приведенными в работах [1–4] рассуждениями.

Согласно вышесказанному, каждый путь является образом кривой без самопересечений, моделирующей траекторию движения устройства генерации следа [2]. Поскольку путь однозначно определяется парой атомарных элементов α_a, $\alpha_b \in A$ и последовательностью связных связок между ними, в работе [3] введено в рассмотрение множество $\Im(\alpha_a, \alpha_b)$ всевозможных путей L_k из α_a в α_b: $\Im(\alpha_a, \alpha_b) = \{L_k\}$, $k \in \{1, 2, ..., K_0\}$, где $\forall k \in \{1, 2, ..., K_0\}$,

$$L_k = L_k(\alpha_a, \alpha_b) = \{(\alpha_h^k, \alpha_{h+1}^k)_{m_h^k}\}_{h=1}^{n_k},\ m_h^k \in \hat{M},\ h = \overline{1, n_k}.$$

Исследования свойств различных путей показали, что для произвольных $\alpha_a, \alpha_b \in A$ пути множества $\Im(\alpha_a, \alpha_b)$, характеризуемые одинаковыми типами образующих их связок, в общем случае моделируют GT-объекты различных изображений знаков, которые не являются одноименными, как показано, в частности, на рис. 2.

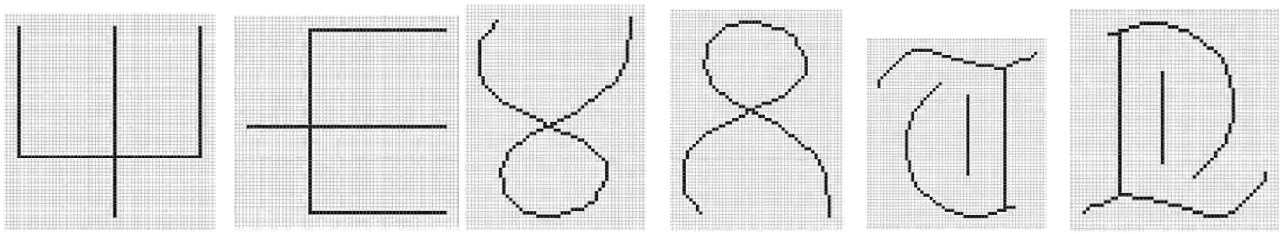


Рис. 2. Множества путей, неразличимых по типам составляющих их связок, моделирующие на множестве АЭ образы GT-объектов различных изображений знаков, не являющихся одноименными

Связки различных типов, формирующие пути, в зависимости от порядка следования, позволяют выявить фрагменты D-знаков, соответствующих исходным изображениям, характеризуемые сохранением локально-глобального направления (ЛГН) [2]. Введенные в [3] меры μ_1 и μ_2 путей, а также результаты исследований их свойств, позволили определить кратчайшие пути между парами произвольных АЭ α_a, $\alpha_b \in A$, характеризуемые сохранением локально-глобального направления в целом [2, 4].

Определение 1. Путь $L(\alpha_a, \alpha_b)$ такой, что

$$\mu_2(L) = \min_{\forall L_k \in \Im_1(\alpha_a, \alpha_b)} \mu_2(L_k), \qquad (2)$$

называется *кратчайшим путем* от α_a к α_b (из α_a в α_b), где $\Im_1(\alpha_a, \alpha_b) \subset \Im(\alpha_a, \alpha_b)$ – множество путей из α_a в α_b таких, что $\forall L_k \in \Im_1(\alpha_a, \alpha_b), k \in K_1$, где $K_1 \subset \{1, 2, ..., K_0\}$ – индексное множество, выполнено $\mu_1(L_k) = \rho(\alpha_a, \alpha_b)$.

В работе [3] показано, что для произвольных АЭ $\alpha_a, \alpha_b \in A$ в общем случае существует множество $\Im_2(\alpha_a, \alpha_b) \subseteq \Im_1(\alpha_a, \alpha_b)$ КП между ними, $\Im_2(\alpha_a, \alpha_b) = \{L_k\}$, $k \in K_2 \subseteq K_1$, где K_2 – индексное множество.

Также обосновано утверждение, согласно которому кратчайшие пути не содержат одновременно горизонтальных и вертикальных связок, а также диагональных связок двух типов, то есть s_1 и s_2 одновременно, или s_3 и s_4 одновременно. Таким образом, КП на множестве АЭ являются моделями принадлежащих GT-объекту некоторого изображения исходного знака образов кривых без самопересечений, характеризуемых сохранением локально-глобального направления движения [4], либо их фрагментов [2].

Заметим также, что, согласно определениям пути и КП [3], для любых АЭ $\alpha_a, \alpha_b \in A$ произвольный путь $L_k \in \Im(\alpha_a, \alpha_b)$, $k \in \{1, 2, ..., K_0\}$, может быть представлен в виде объединения конечного числа кратчайших путей.

Приведенный в [3] критерий принадлежности произвольного пути множеству кратчайших в совокупности с доказанными утверждениями позволяют осуществлять автоматическую декомпозицию произвольного D-знака, предполагающую выявление фрагментов образов GT-объектов, которые соответствуют кривым модели исходного изображения знака [2], и характеризуются сохранением ЛГН.

Конструктивизм определения указанных фрагментов образов GT-объектов на множестве АЭ позволяет осуществлять их автоматическое выявление без априорного задания элементов открытого алфавита и множеств непроизводных элементов.

Таким образом, КП является искомой элементарной составляющей моделей образов GT-объектов на множестве АЭ, не предполагающей априорного задания множеств непроизводных элементов, эталонов, примитивов, последовательностей морфологических операций, которая позволяет в силу конструктивизма своего определения и имеющихся теоретических предпосылок [1–4] автоматически формировать модели анализируемых образов бинарных РЦИ знаков открытых алфавитов и их GT-объектов в терминах свойств дискретного множества АЭ.

Выбор КП в качестве элементарной составляющей модели D-знака предполагает возможность автоматического формирования математической модели образа GT-объекта произвольного бинарного РЦИ аналогично структурным методам [6]. Разница с указанными методами заключается в том, что, в отличие от непроизводных элементов, КП задается конструктивно безотносительно к объектам обучающей выборки и их эвристически-субъективному анализу на этапе проектирования системы распознавания, и его автоматическое выявление не предполагает необходимости применения мер близости, пороговых констант [6] и прочих способов [5], снижающих, как показано в [1], качество распознавания.

На рис. 3. приведены примеры результатов автоматизированного выявления на образах GT-объектов множеств КП, каждый из которых характеризуется на множестве АЭ сохранением ЛГН [2].

Заметим, что наличие в КП между парой фиксированных АЭ связок установленных типов в общем случае предполагает наличие различий между КП в рамках одного множества $\Im_2(\alpha_a, \alpha_b)$, обусловленных различными порядками следования связок, формирующих различные кратчайшие пути. Данный факт подтверждается, в частности, примерами, приведенными на рис. 4.

Указанные различия между КП, принадлежащими одному множеству $\Im_2(\alpha_a, \alpha_b)$, в общем случае предопределяют проблематичность именования образов GT-объектов как совокупностей кратчайших путей, сформированных одинаковыми типами связок.

На рис. 5 приведены примеры множеств КП, выявленных на образах GT-объектов, неразличимых по типам составляющих связок с аналогичными путями, выявленными на образах GT-объектов.

Таким образом, КП, являясь конструктивно определенным на множестве АЭ универсальным аналогом непроизводного элемента, при его использовании в качестве элементарной составляющей автоматически формируемой ММ образа GT-объекта бинарного РЦИ знака произвольного алфавита, предполагает необходимость учитывать его свойства, отражающие в частности, последовательности типов составляющих его связок. То есть кратчайшие пути, изображенные на рис. 4, целесообразно различать между собой.

2 ОПРЕДЕЛЕНИЕ КАЧЕСТВЕННЫХ ХАРАКТЕРИСТИК КРАТЧАЙШИХ ПУТЕЙ

В соответствии с проведенными рассуждениями, для того, чтобы различать принадлежащие одному множеству $\Im_2(\alpha_a, \alpha_b)$ пути, порядки следования связок в которых различны, определим понятия выпуклых вверх, вниз, влево и вправо КП. Указанные характеристики в некотором смысле аналогичны понятиям выпуклости и вогнутости, определенных для кривых в терминах свойств всюду плотных множеств и позволяющих, в частности, учитывать локальные свойства соответствующих объектов, моделирующих изображения знаков, например, в признаковых [5] и структурных [6] подходах к распознаванию.

Определение 2. Путь $L_k \in \Im_2(\alpha_a, \alpha_b)$, $L_k = L_k(\alpha_a, \alpha_b) = \{(\alpha_h^k, \alpha_{h+1}^k)_{m_h^k}\}_{h=1}^{n_k}$, $m_h^k \in \hat{M}$, $h = \overline{1, n_k}$, $k \in K_2$, называется *верхней границей* множества кратчайших путей, если $\forall L_l \in \Im_2(\alpha_a, \alpha_b)$, $L_l = L_l(\alpha_a, \alpha_b) = \{(\alpha_r^l, \alpha_{r+1}^l)_{m_r^l}\}_{r=1}^{n_k}$, $m_r^l \in \hat{M}$, $r = \overline{1, n_k}$, $l \in K_2$, и $\forall \alpha_h = \alpha(i_h, j_h) \in \Lambda(L_k)$, $\forall \alpha_r = \alpha(i_r, j_r) \in \Lambda(L_l)$,

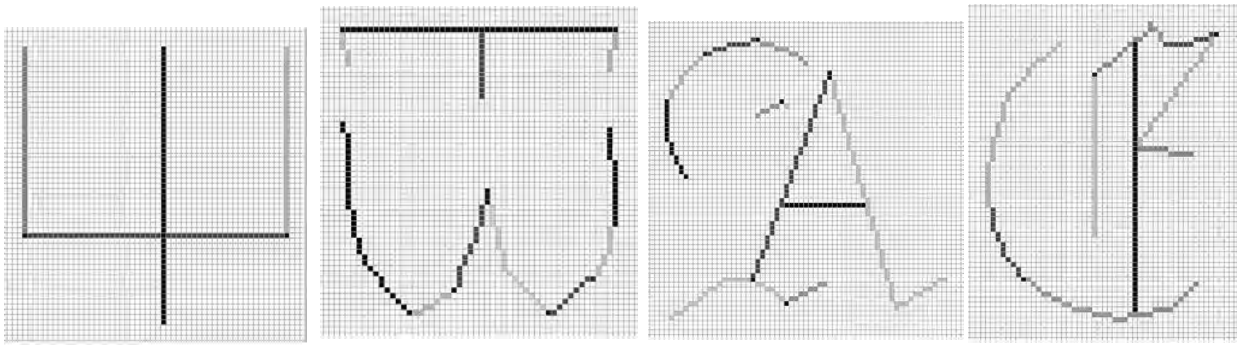


Рис. 3. Множества КП, выявленных на образах GT-объектов в автоматизированном режиме

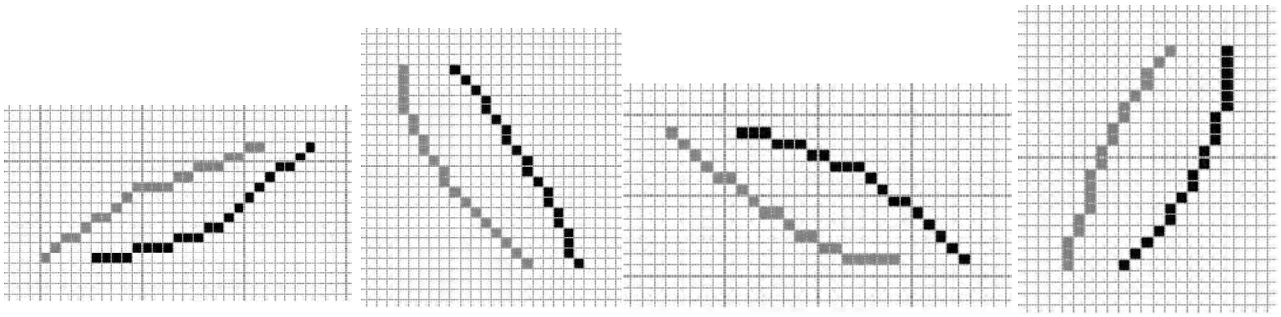


Рис. 4. Примеры КП между парами фиксированных АЭ, моделирующих локальные изменения направлений движений РЧ УФС

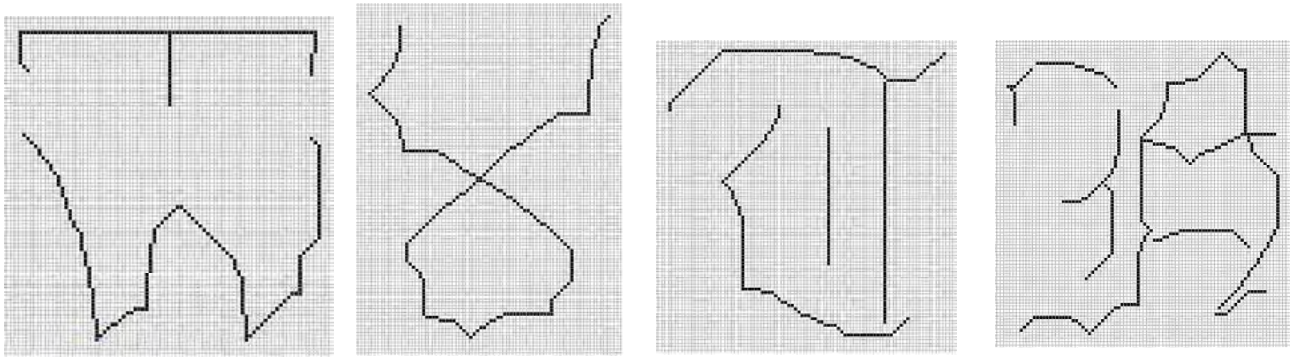


Рис. 5. Образы GT-объектов, именование которых проблематично

$h, r = \overline{1, n_k + 1}$, где L_l – произвольный КП из α_a в α_b, выполнено:

$$i_h \le i_r,\ j_h \le j_r, \text{ в случае, когда } m_h^k,\ m_r^l \in \{1, 4\},\ h, r = \overline{1, n_k}$$

или

$$i_h \le i_r,\ j_h \ge j_r, \text{ в случае, когда } m_h^k,\ m_r^l \in \{1, 3\},\ h, r = \overline{1, n_k}. \quad (3)$$

Определение 3. Путь $L_k \in \Im_2(\alpha_a, \alpha_b)$, $L_k = L_k(\alpha_a, \alpha_b) = \{(\alpha_h^k, \alpha_{h+1}^k)_{m_h^k}\}_{h=1}^{n_k},\ m_h^k \in \hat{M},\ h = \overline{1, n_k}$, $k \in K_2$, называется *левой границей* множества кратчайших путей, если $\forall L_l \in \Im_2(\alpha_a, \alpha_b)$, $L_l = L_l(\alpha_a, \alpha_b) = \{(\alpha_r^l, \alpha_{r+1}^l)_{m_r^l}\}_{r=1}^{n_k},\ m_r^l \in \hat{M},\ r = \overline{1, n_k}$, $l \in K_2$, и $\forall \alpha_h = \alpha(i_h, j_h) \in \Lambda(L_k)$, $\forall \alpha_r = \alpha(i_r, j_r) \in \Lambda(L_l)$, $h, r = \overline{1, n_k + 1}$, где L_l – произвольный КП из α_a в α_b, выполнено:

$$i_h \le i_r,\ j_h \le j_r, \text{ в случае, когда } m_h^k,\ m_r^l \in \{2, 4\},\ h, r = \overline{1, n_k}, \text{ или}$$

$$i_h \ge i_r,\ j_h \le j_r, \text{ в случае, когда } m_h^k,\ m_r^l \in \{2, 3\},\ h, r = \overline{1, n_k}. \quad (4)$$

Примеры верхних и левых границ различных множеств кратчайших путей приведены на рис. 6.

Аналогично определим нижнюю и правую границы множества КП.

Определение 4. Путь $L_k \in \Im_2(\alpha_a, \alpha_b)$, $L_k = L_k(\alpha_a, \alpha_b) = \{(\alpha_h^k, \alpha_{h+1}^k)_{m_h^k}\}_{h=1}^{n_k},\ m_h^k \in \hat{M},\ h = \overline{1, n_k}$, $k \in K_2$, называется *нижней границей* множества кратчайших путей, если $\forall L_l \in \Im_2(\alpha_a, \alpha_b)$, $L_l = L_l(\alpha_a, \alpha_b) = \{(\alpha_r^l, \alpha_{r+1}^l)_{m_r^l}\}_{r=1}^{n_k},\ m_r^l \in \hat{M},\ r = \overline{1, n_k}$, $l \in K_2$, и $\forall \alpha_h = \alpha(i_h, j_h) \in \Lambda(L_k)$, $\forall \alpha_r = \alpha(i_r, j_r) \in \Lambda(L_l)$, $h, r = \overline{1, n_k + 1}$, где L_l – произвольный КП из α_a в α_b, выполнено:

$$i_h \ge i_r,\ j_h \ge j_r, \text{ в случае, когда } m_h^k,\ m_r^l \in \{1, 4\},\ h, r = \overline{1, n_k}, \text{ или}$$

$$i_h \ge i_r,\ j_h \le j_r, \text{ в случае, когда } m_h^k,\ m_r^l \in \{1, 3\},\ h, r = \overline{1, n_k}. \quad (5)$$

Определение 5. Путь $L_k \in \Im_2(\alpha_a, \alpha_b)$, $L_k = L_k(\alpha_a, \alpha_b) = \{(\alpha_h^k, \alpha_{h+1}^k)_{m_h^k}\}_{h=1}^{n_k},\ m_h^k \in \hat{M},\ h = \overline{1, n_k}$, $k \in K_2$, называется *правой границей* множества кратчайших путей, если $\forall L_l \in \Im_2(\alpha_a, \alpha_b)$, $L_l = L_l(\alpha_a, \alpha_b) = \{(\alpha_r^l, \alpha_{r+1}^l)_{m_r^l}\}_{r=1}^{n_k},\ m_r^l \in \hat{M},\ r = \overline{1, n_k}$, $l \in K_2$, и $\forall \alpha_h = \alpha(i_h, j_h) \in \Lambda(L_k)$, $\forall \alpha_r = \alpha(i_r, j_r) \in \Lambda(L_l)$, $h, r = \overline{1, n_k + 1}$, где L_l – произвольный КП из α_a в α_b, выполнено:

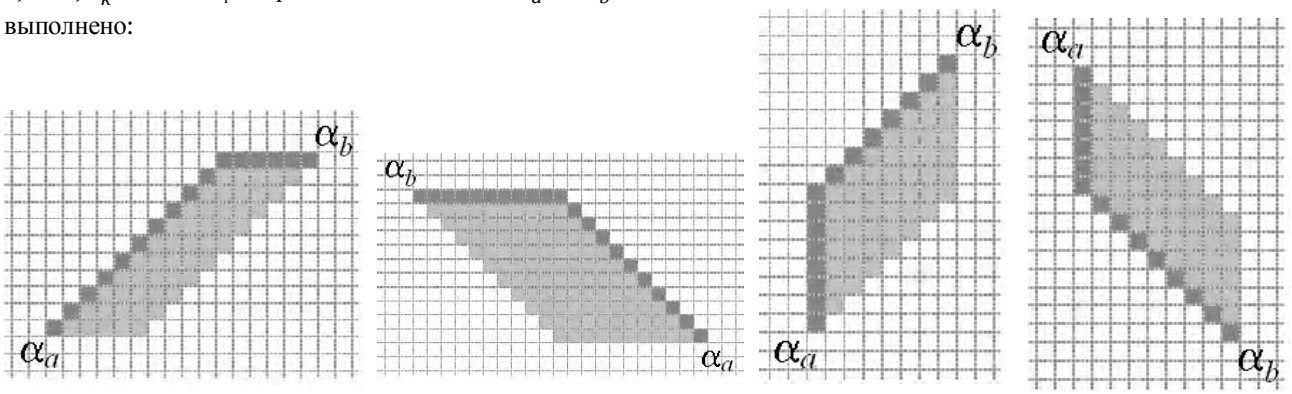
$$i_h \ge i_r,\ j_h \ge j_r, \text{ в случае, когда } m_h^k,\ m_r^l \in \{2, 4\},\ h, r = \overline{1, n_k}, \text{ или}$$

$$i_h \le i_r,\ j_h \ge j_r, \text{ в случае, когда } m_h^k,\ m_r^l \in \{2, 3\},\ h, r = \overline{1, n_k}. \quad (6)$$

На рис. 7 приведены примеры нижних и правых границ множеств КП.

На основании предложенных определений выделим в различных множествах КП $\Im_2(\alpha_a, \alpha_b)$ подмножества путей, характеризуемых, согласно проведенным рассуждениям, верхней, нижней, левой либо правой ориентацией, что позволит при автоматическом моделировании и именовании на множестве АЭ образов бинарных РЦИ знаков открытых алфавитов более точно учитывать особенности сохраняющих ЛГН фрагментов образов GT-объектов [2].

Определение 6. Путь $L_k \in \Im_2(\alpha_a, \alpha_b)$, $L_k = L_k(\alpha_a, \alpha_b) = \{(\alpha_h^k, \alpha_{h+1}^k)_{m_h^k}\}_{h=1}^{n_k},\ m_h^k \in \hat{M},\ h = \overline{1, n_k}$, $k \in K_2$, называется *ориентированным вверх*, если $\forall \alpha_h^k = \alpha(i_h^k, j_h^k) \in \Lambda(L_k),\ h = \overline{2, n_k}$, выполнено следую-



а) *б)* *в)* *г)*

Рис. 6. Верхние (*а*, *б*) и левые (*в*, *г*) границы множеств КП

щее условие:

$$j_h^k = j_h^t = j_h^l,\ i_h^t \le i_h^k \le i_h^t + \left[\frac{i_h^l - i_h^t - 1}{2}\right],$$

при $m_h^k \in \{1, 3\}$ либо $m_h^k \in \{1, 4\}$,

где $i_h^t,\ j_h^t,\ i_h^l,\ j_h^l$ – индексы АЭ $\alpha_h^t = \alpha(i_h^t,\ j_h^t) \in \Lambda(L_t)$, $\alpha_h^l = \alpha(i_h^l,\ j_h^l) \in \Lambda(L_l)$, $L_t, L_l \in \Im_2(\alpha_a, \alpha_b)$, $t, l \in K_2$, – соответственно верхняя и нижняя границы множества $\Im_2(\alpha_a, \alpha_b)$.

На рис. 8, *а*), *б*) приведены примеры множеств АЭ, формирующих связки ориентированных вверх КП различных множеств $\Im_2(\alpha_a, \alpha_b)$ в соответствии с определением 6.

Определение 7. Путь $L_k \in \Im_2(\alpha_a, \alpha_b)$, $L_k = L_k(\alpha_a, \alpha_b) = \{(\alpha_h^k, \alpha_{h+1}^k)_{m_h^k}\}_{h=1}^{n_k},\ m_h^k \in \hat{M},\ h = \overline{1, n_k}$, $k \in K_2$, называется *ориентированным влево*, если $\forall \alpha_h^k = \alpha(i_h^k,\ j_h^k) \in \Lambda(L_k)$, $h = \overline{2, n_k}$, выполнено:

$$i_h^k = i_h^t = i_h^l,\ j_h^t \le j_h^k \le j_h^t + \left[\frac{j_h^l - j_h^t - 1}{2}\right], \text{ при } m_h^k \in \{2, 4\},$$

либо $m_h^k \in \{2, 3\}$,

где $i_h^t,\ j_h^t,\ i_h^l,\ j_h^l$ – индексы АЭ $\alpha_h^t = \alpha(i_h^t,\ j_h^t) \in \Lambda(L_t)$, $\alpha_h^l = \alpha(i_h^l,\ j_h^l) \in \Lambda(L_l)$, $L_t, L_l \in \Im_2(\alpha_a, \alpha_b)$, $t, l \in K_2$, – соответственно левая и правая границы множества $\Im_2(\alpha_a, \alpha_b)$.

Рис. 8, в), г) содержат примеры множеств АЭ, формирующих связки КП, ориентированных влево согласно определению 7, для различных множеств $\Im_2(\alpha_a, \alpha_b)$.

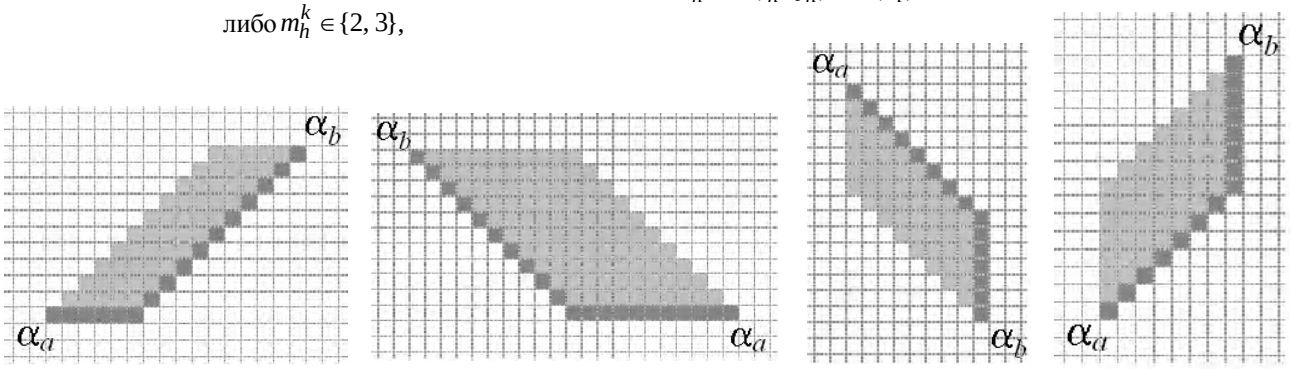
Аналогично ориентированным вверх и влево, для различных множеств $\Im_2(\alpha_a, \alpha_b)$ определим ориентированные вниз и вправо КП.

Определение 8. Путь $L_k \in \Im_2(\alpha_a, \alpha_b)$, $L_k = L_k(\alpha_a, \alpha_b) = \{(\alpha_h^k, \alpha_{h+1}^k)_{m_h^k}\}_{h=1}^{n_k},\ m_h^k \in \hat{M},\ h = \overline{1, n_k}$, $k \in K_2$, называется *ориентированным вниз*, если $\forall \alpha_h^k = \alpha(i_h^k,\ j_h^k) \in \Lambda(L_k)$, $h = \overline{2, n_k}$, выполнено следующее условие:

$$j_h^k = j_h^t = j_h^l,\ i_h^l - \left[\frac{i_h^l - i_h^t - 1}{2}\right] \le i_h^k \le i_h^l, \text{ при } m_h^k \in \{1, 3\}$$

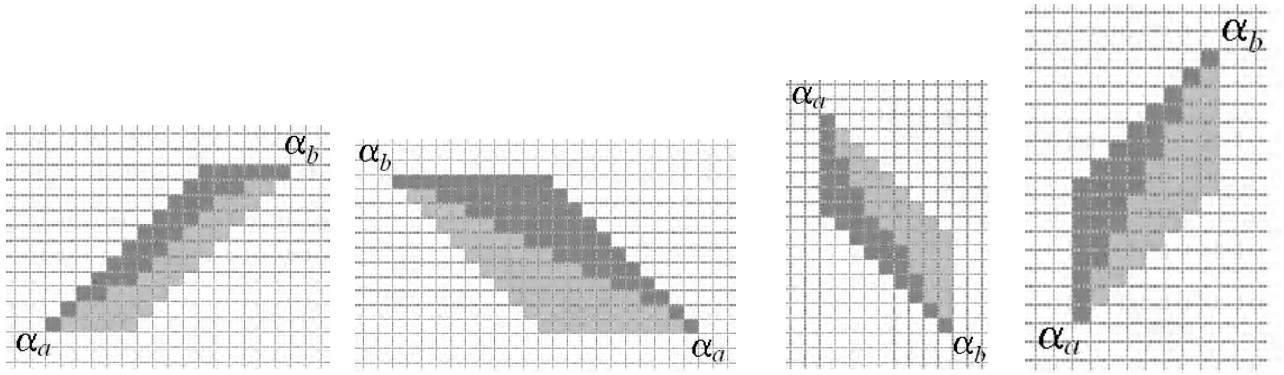
либо $m_h^k \in \{1, 4\}$,

где $i_h^t,\ j_h^t,\ i_h^l,\ j_h^l$ – индексы АЭ $\alpha_h^t = \alpha(i_h^t,\ j_h^t) \in \Lambda(L_t)$, $\alpha_h^l = \alpha(i_h^l,\ j_h^l) \in \Lambda(L_l)$, $L_t, L_l \in \Im_2(\alpha_a, \alpha_b)$, $t, l \in K_2$, –



а) *б)* *в)* *г)*

Рис. 7. Нижние (*а*, *б*) и правые (*в*, *г*) границы множеств КП



а) *б)* *в)* *г)*

Рис. 8. Множества АЭ, формирующие связки ориентированных вверх (*а*, *б*) и влево (*в*, *г*) КП множеств $\Im_2(\alpha_a, \alpha_b)$ (выделены темным)

соответственно верхняя и нижняя границы множества $\Im_2(\alpha_a, \alpha_b)$.

Определение 9. Путь $L_k \in \Im_2(\alpha_a, \alpha_b)$, $L_k = L_k(\alpha_a, \alpha_b) = \{(\alpha_h^k, \alpha_{h+1}^k)_{m_h^k}\}_{h=1}^{n_k}$, $m_h^k \in \hat{M}$, $h = \overline{1, n_k}$, $k \in K_2$, называется *ориентированным вниз*, если $\forall \alpha_h^k = \alpha(i_h^k, j_h^k) \in \Lambda(L_k)$, $h = \overline{2, n_k}$, выполнено следующее условие:

$$j_h^k = j_h^t = j_h^l,\ i_h^l - \left[\frac{i_h^l - i_h^t - 1}{2}\right] \le i_h^k \le i_h^l,$$

при $m_h^k \in \{1, 3\}$ либо $m_h^k \in \{1, 4\}$,

где i_h^t, j_h^t, i_h^l, j_h^l – индексы АЭ $\alpha_h^t = \alpha(i_h^t, j_h^t) \in \Lambda(L_t)$, $\alpha_h^l = \alpha(i_h^l, j_h^l) \in \Lambda(L_l)$, $L_t, L_l \in \Im_2(\alpha_a, \alpha_b)$, $t, l \in K_2$ – соответственно верхняя и нижняя границы множества $\Im_2(\alpha_a, \alpha_b)$.

Примеры множеств АЭ, формирующих, в соответствии с определениями 8, 9, связки ориентированных вниз и вправо КП из различных множеств $\Im_2(\alpha_a, \alpha_b)$, приведены на рис. 9, *а*), *б*) и 9, *в*), *г*) соответственно.

В случае, если произвольный путь $L_k \in \Im_2(\alpha_a, \alpha_b)$, $k \in K_2$, не является ориентированным вверх, вниз, влево либо вправо согласно определениям 6–9, он может быть представлен в виде объединения конечного числа ориентированных КП.

Если начальные и конечные АЭ КП некоторого множества $\Im_2(\alpha_a, \alpha_b)$ таковы, что $i_a = i_b$ или $j_a = j_b$, но при этом $i_a \ne i_b$ и $j_a \ne j_b$ одновременно, или $|i_a - i_b| = |j_a - j_b|$, то для фиксированных значений индексов АЭ $\alpha_a, \alpha_b \in A$, удовлетворяющих указанным условиям, существует единственный КП между ними, который состоит из связок только одного типа, при этом верхняя и нижняя либо левая и правая границы множеств КП совпадают (рис. 10).

В данном случае соответствующие образы кривых как элементы GT-объекта, или их фрагменты, характеризуемые сохранением ЛГН, на множестве АЭ моделируются D-отрезками [3], являющимися частными случаями КП. При этом, согласно [1], вариативность в выборе элементарной составляющей модели образа GT-объекта минимальна, следовательно, определение ориентированности кратчайшего пути не является целесообразным, и автоматическое моделирование осуществляется непосредственно к выявленным D-отрезкам.

ЗАКЛЮЧЕНИЕ

Выбор кратчайшего пути в качестве элементарной составляющей модели образа GT-объекта изображения произвольного знака открытого алфавита позволяет, в частности, эффективно осуществлять автоматическую

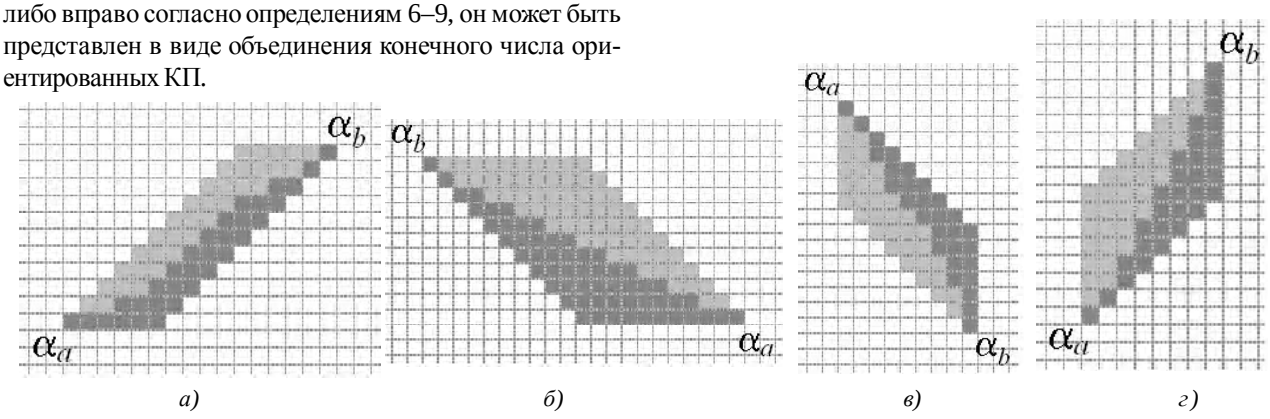


Рис. 9. Множества АЭ, формирующие связки ориентированных вниз (*а*, *б*) и вправо (*в*, *г*) КП множеств $\Im_2(\alpha_a, \alpha_b)$ (выделены темным)

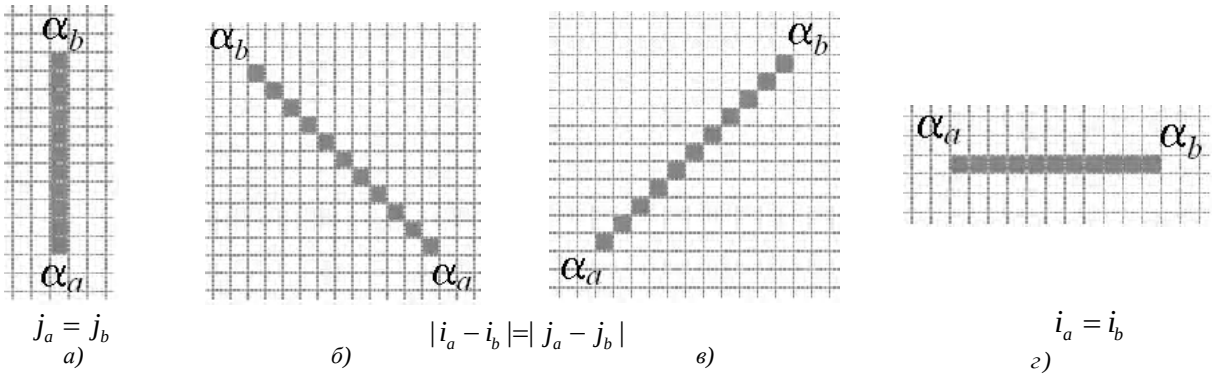


Рис. 10. Множества КП, состоящие из единственного элемента, являющегося одновременно левой и правой (*а*–*в*) либо верхней и нижней (*б*–*г*) границами множеств $\Im_2(\alpha_a, \alpha_b)$

декомпозицию D-знаков [3]. Определения ориентированных вверх, вниз влево и вправо КП позволяют при формировании моделей анализируемых D-знаков учитывать такую характеристику кривых, как выпуклость, традиционно используемую, согласно [5, 6], при моделировании исходных изображений знаков.

Проведенные в настоящей работе рассуждения и введенные определения позволяют перейти к рассмотрению проблемы взаимного расположения КП в моделируемых D-знаках, что позволит учитывать относительные характеристики автоматически выявляемых в процессе декомпозиции элементарных составляющих образов GT-объектов, представимых в виде множества путей [4], соответствующих порождающим траекториям исходных изображений знаков, сгенерированных в соответствии с изложенной в [2] гипотезой.

СПИСОК ЛИТЕРАТУРЫ

1. *Шевцов, Д. В.* Обоснование перспективных направлений при проектировании систем автоматизированной обработки видеоинформации [Текст] / Д. В. Шевцов // Вісник ХНТУ. – 2009. – № 1 (34). – С. 231–240.
2. *Мельник, А.-В. В.* Моделируемость GT-объекта на дискретном множестве атомарных элементов [Текст] / А.-В. В. Мельник, С. В. Мышко, Д. В. Шевцов // Вісник ХНТУ. –2008. – № 33. – С. 112–118.
3. *Мышко, С. В.* Автоматическая декомпозиция изображений при их опознавании в системах технического зрения роботов [Текст] / С. В. Мышко, Д. В. Шевцов // Наукові праці Донецького державного технічного університету. Серія : Обчислювальна техніка та автоматизація. Випуск 38. – Донецьк : РВА ДонДТУ, 2002. – С. 216–222.
4. *Вайсруб, Н. В.* Применение способа формирования GT-объектов и способа автоматического моделирования изображений при проектировании систем технического зрения [Текст] / Н. В. Вайсруб, А.-В. В. Мельник, Д. В. Шевцов // Інформаційно-керуючи системи на залізничному транспорті. – 2009. – № 2. – С. 55–59.
5. *Гонсалес, Р.* Принципы распознавания образов [Текст] / Р. Гонсалес, Дж. Ту ; пер. с англ. – М. : Мир, 1978. – 416 с.
6. *Фу, К.* Структурные методы в распознавании образов [Текст] / К. Фу ; пер. с англ. – М. : Мир, 1977. – 319 с.

Стаття надійшла до редакції 28.08.2013.
Після доробки 24.10.2013.

Шевцов Д. В.
Канд. техн. наук, доцент, Донецький національний університет, Україна

ВИЗНАЧЕННЯ ОРІЄНТАЦІЇ ЕЛЕМЕНТАРНИХ СКЛАДОВИХ МОДЕЛЕЙ ЗНАКІВ, ЯКІ ПІДЛЯГАЮТЬ АВТОМАТИЧНОМУ ІМЕНУВАННЮ НА МНОЖИНІ АТОМАРНИХ ЕЛЕМЕНТІВ

Автоматичне моделювання, іменування та опізнавання знаків цифрових бінарних зображень довільної природи є актуальною науково-практичною задачею, яка знаходить своє використання в багатьох галузях впровадження інформаційних технологій, зокрема, при обробці й аналізі електронних документів. Стаття присвячена конструктивному визначенню структурних складових зображень, які модулюються після скелетизації при попередній обробці, та їх властивостей, які нададуть можливості здійснювати автоматичне виявлення вказаних об'єктів з метою наступного формування описів знаків, що іменуються та розпізнаються на дискретній множині атомарних елементів.

Ключові слова: автоматичне моделювання, цифрові бінарні зображення, іменування, розпізнавання, декомпозиція, найкоротший шлях.

Shevtsov D. V.
Ph.D., associate Professor, Donetsk National University, Ukraine

DETERMINING THE ORIENTATION OF THE ELEMENTARY COMPONENTS OF CHARACTERS MODELS, SUBJECT TO AUTOMATIC NAMING ON THE SET OF ATOMIC ELEMENTS

Arbitrary nature digital binary images signs automatic modeling, naming and recognition are actual scientific and practical problems, which are existing in many areas of information technologies application scopes, particularly in electronic documents processing and analyzing. This article is devoted to a constructive definition of images structure components, that are modeled after skeletonization during preprocessing, and their properties, which will give an opportunity to perform automatic detection of described objects in order to form description of future signs, which will be named and recognized on atomic elements discrete set

Keywords: automatic modeling, digital binary images, naming, recognition, decomposition of the shortest path.

REFERENCES

1. Shevczov D. V. Obosnovanie perspektivny'h napravlenij pri proektirovanii sistem avtomatizirovannoj obrabotki videoinformacii, *Visnyk HNTU*, 2009, No.1 (34), pp. 231–240.
2. Mel'nik A.-V. V., My'shko S. V., Shevczov D. V. Modeliruemoct' GT-ob''ekta na diskretnom mnozhestve atomarny'h e'lementov, *Visnik HNTU*, 2008, No. 33, pp. 112–118.
3. My'shko S. V., Shevczov D. V. Avtomaticheskaya dekompoziciya izobrazhenij pri ih opoznavanii v sistemah texnicheskogo zreniya robotov, *Naukovi pratsi Donets'kogo derzhavnogo tehnichnogo universitetu. Seriia: Obchysliuval'na tehnika ta avtomatyzatsiia, Vypusk 38*, Donets'k, RVA DonDTU, 2002, pp. 216–222.
4. Vajsrub N. V., Mel'nik A.-V. V., My'shko S. V., Shevczov D. V. Primenenie sposoba formirovaniya GT-ob''ektov i sposoba avtomaticheskogo modelirovaniya izobrazhenij pri proektirovanii sistem tehnicheskogo zreniya, *Informatsiino-keruiuchi systemy na zaliznychnomu transporti*, 2009, No. 2, pp. 55–59.
5. Gonsales R., Tu Dzh Principy' raspoznavaniya obrazov, Moscow, Mir, 1978, 416 p.
6. Fu K. Strukturny'e metody' v raspoznavanii obrazov, Moscow, Mir, 1977, 319 p.

НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

НЕЙРОИНФОРМАТИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

NEUROINFORMATICS AND INTELLIGENT SYSTEMS

UDC 681.5.015:629.7.05

Firsov S. N.[1], Reznikova O. V.[2]

[1]Ph. D., Associate Professor, National Aerospace University named after M. E. Zhukovsky «KhAI», Ukraine, E-mail: sn.firsof@gmail.com

[2]Assistant, National Aerospace University named after M. E. Zhukovsky «KhAI», Ukraine

FAULT TOLERANCE OF SPACECRAFT ORIENTATION AND STABILIZATION SYSTEM

Formed the structure of the orientation and stabilization system for ensuring its active fault tolerance. Parameterized types of failures of the system functional elements, characterized by changes in the properties of the conversion elements. Received diagnostic functional models for solving problems of detection, search the place and establish a class of failure for system. Developed models and methods of systematic approach to fault tolerance in the direction of the effective use of the signal, parametric and structural redundancies and selection of parrying tools. Performed experimental researches of the sample model of the automatic attitude control and stabilization the properties of active fault-tolerance in emergency modes of operation.

Keywords: stabilization and orientation spacecraft system, fault tolerance, reaction wheel, angular velocity sensor.

Analysis of priority and future space programs shows that the actual is the establishment and maintenance of small satellites with operation period over 15 years [1]. The solution of this problem involves various difficulties, including the fault of spacecraft blocks and systems. One promising avenue is to create a fault-tolerant spacecraft orientation and stabilization systems.

Small spacecraft conditions of use involves a high precision working off program reversals with high accuracy stabilization of the angular position. These conditions with acceptable energy and mass-dimensional characteristics are able to ensure reaction wheels (RW) as the executive bodies [2]. An increase in resources of such systems is achieved by applying the minimally redundant system schemes of RW [3].

Since any deviation of the spacecraft system parameters can lead to a deviation from stated rigid requirements for their work, it is impossible to use traditional approaches to fault tolerance associated with redundancy of its con-stouten elements [4]. Therefore, necessary to adopt the application methods and models associated with providing of an active fault tolerance.

Significant contribution to the development of models, methods and means of active fault tolerance was made by such scientists as R. Isermann, R. Patton, P. Frank, R. Beard, V. Y. Rutkowski, A. V. Mozgalevskaya, I. V. Kuzmin, B. I. Dotsenko, J. E. Eisenberg, D. V. Lebedev, A. S. Kulik, V. S. Blintsov, L. G. Raskin. However, the proposed models, methods and tools to ensure the active fault-tolerance are predominantly fragmented and they do not reflect the dynamics of the control processes, do not con-sider the possibilities of diagnosing the operational state of the functional elements, as well as fault parrying through the effective use of existing and the introduction of additional redundancies. Consequently, the development of diagnosing models and methods, as well as failure parrying to provide an active fault-tolerance of the spacecraft attitude determination and control in real-time is an important scientific and applied problem.

PROBLEM FORMULATION

To achieve this purpose the authors have solved the following problems:

1. Determined the structure of the spacecraft attitude determination and control system, set of its functional elements failures types and developed diagnostic models that connect direct and indirect features.

2. The models and methods for in-depth diagnosis of the spacecraft orientation and stabilization system with depth to failure type are developed.

3. The models and methods for parrying failures in the spacecraft attitude determination and control system through effective use of the signal, parametric and structural redundancies are designed.

4. The experimental procedures for the testing of active fault-tolerance model sample of spacecraft attitude determination and control system are held. Also evaluated its comprehensive indicator of the quality and fault tolerance level.

DETERMINATION OF THE SPACECRAFT ORIENTATION AND STABILIZATION SYSTEM STRUCTURE

Consider the generalized block diagram of the perspective spacecraft orientation control system (Fig. 1).

A typical control system of the perspective spacecraft includes in the structure the spacecraft as a control object, RW for generating control moments, complex of angular velocity gauges (CAVG) for measuring angular velocity, data spacecraft subsystem for generating control signals applied to all blocks of the system. To orient the spacecraft and CAVG data correction at the system uses celestial navigation system (CNS) and combined orientation sensor (COS). Electromagnetic drives driven by signals from the magnetometers installed on spacecraft in order to unload RW.

The initial configurations of accommodating RW on the spacecraft with minimum redundancy were selected following variants [5]:

– scheme provided for NASA standards within the project multipurpose modular platform MMS, shown in Fig. 2, *a*;

– the scheme of the executive drives installation of the company General Electric, with RW kinetic moments directed from the middle of the cube (Fig. 2, *b*).

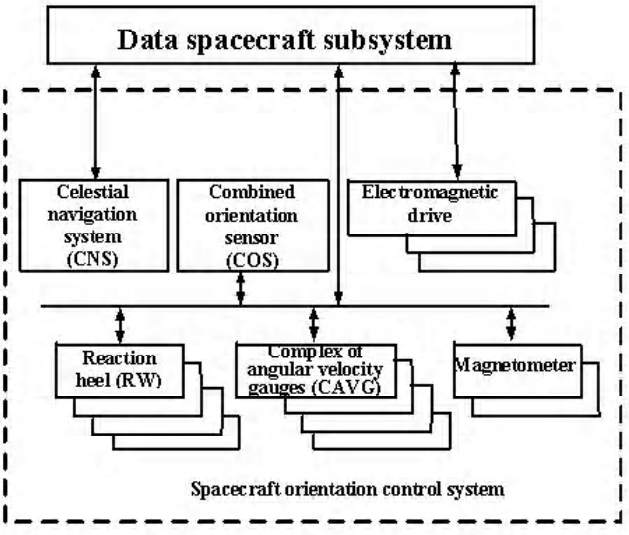


Fig. 1. Structural diagram of the perspective spacecraft control system

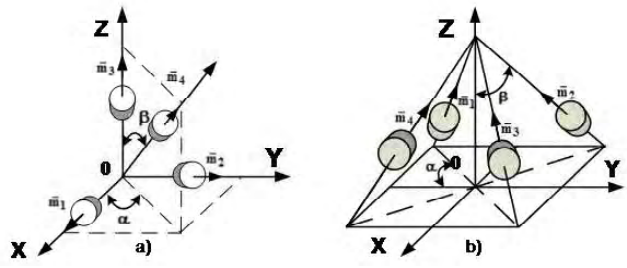


Fig. 2. Minimally redundant RW installation schemes on the spacecraft

It is known that when comparing different configurations that satisfy the technical requirements, preference is given to those who provide the field of formation control torque to the ellipsoid of the following form:

$$\left(\frac{M_x}{I_x\dot{\omega}_x}\right)^2+\left(\frac{M_y}{I_y\dot{\omega}_y}\right)^2+\left(\frac{M_z}{I_z\dot{\omega}_z}\right)^2=1. \qquad (1)$$

where M_x, M_y, M_z are projections of the total required reactive RW torque on the axis of the coordinate system associated (CSA); I_x, I_y, I_z are spacecraft axial moments of inertia; $\dot{\omega}_x, \dot{\omega}_y, \dot{\omega}_z$ are projections of the desired angular acceleration on the axis of the spacecraft CSA.

Analysis of the shape of the area under consideration for the minimally redundant set of schemes for the cases of the use of all four RW and a failure of one of them showed that in both modes shape of the control torque for the pyramid scheme is closer to the required one. Thus, this scheme is preferable. Also identified the optimal angles of installation RW by analyzing projections of the trace of the correlation matrix of the errors (AS Kulik Saturday AM, Reznikov OV).

The result was that the optimal setting angles for the selected scheme are $\alpha = 45°$ and $\beta = 54,7°$.

At the block diagram of the spacecraft control system also contains a gauges block, changing the dynamics of the system. As the gauges on this object were selected angular velocity sensor (AVS), combined in a block. In this case the AVS scheme of installation that enables the CAVG deep diagnosing with depth to failure mode has been selected the following way: setting of three main AVS vector directed along the axes of the CSA and the installation of additional AVS vector collinear to them (Fig. 3).

The analysis of such a setting scheme performed in the works [6–8], who showed that it provides a CAVG diagnosable with depth to failure mode.

SPACECRAFT DIAGNOSING SYSTEM DEVELOPING

In this paper, the spacecraft system of orientation and stabilization is considered as a combination of automatic control devices (ACD) and the object of automatic control (OAC) (Fig. 4), and diagnosing problem reduces to determining the functional state of the OAC, including the executive bodies of the spacecraft and the feedback sensors.

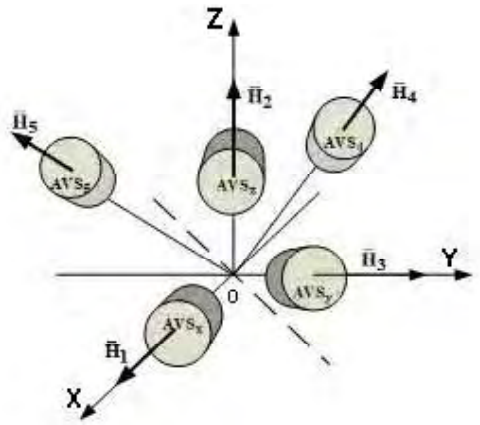


Fig. 3. The installation scheme of AVS on the spacecraft

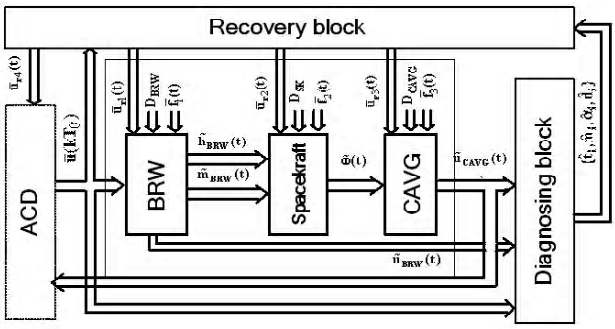


Fig. 4. Functional scheme of spacecraft orientation and stabilization system with active fault tollerance

Many types of failure for the spacecraft orientation and stabilization system D_{SOSS} can be separated into many types of failures of the block of reaction wheels (BRW) D_{BRW}, spacecraft D_{spa} and complex of angular velocity gauges D_{CAVG}.

In this paper, we do not consider the set of ACD failures, which is the microcontroller. Fault tolerance of controllers described in many papers, the current trend – is to use a spacecraft centralized controllers with a three-or five-fold redundancy of the main computer unit. Therefore, consider that the failures of the ACD are parried by switch of reserve units.

Many types of BRW failures consists of sets of RW failure modes, and each RW incorporates the PA, contactless DC motor and tacho, respectively, considered to be 60 kinds of BRW failures.

Many types of failures for CAVG contains multiple failure modes each of the AVS separately, total 30 kinds of failures.

In this paper, we also consider the many types of failures related to the change the moments of inertia of the spacecraft $D_{spa_{IM}}$.

One cause of this failure is the spacecraft incomplete or asymmetric disclosure of solar panels. As shown in various reports incomplete disclosure may occur due to a structural defect in the mechanism of the disclosure, a surface tear of solar cell, mounting damage one of the solar panels at start, etc.

When used as an unloading the jet engine, is possible asymmetrical consumption of fuel and as a result a change in one of the spacecraft moment of inertia.

Many types of spacecraft failures also includes a variety of failure modes associated with the inaccuracy of the external disturbances mathematical model $D_{spa_{dist}}$.

Diagnosing these failure modes will be carried out on two levels: the block and the system. There are only three levels of the hierarchy of diagnostic software (Fig. 5).

First, the lower level is the block level. Here formed diagnostic software (DS) for units of sensors, actuators, calculators not autonomous, but on the basis of their functioning conditions in a closed loop.

The second level is system. Diagnostic software is developed for the entire closed-loop control system of orientation and stabilization. Here is diagnosed how the system runs its functions under the conditions of space flight. The third level of the hierarchy is oversystem. At this level formed diagnostic software for a space mission as a whole.

Block level will form the diagnostic software for CAVG and BRW based on the conditions of their operation in the contour of spacecraft orientation and stabilization system. It allows to provide orientation and stabilization system by blocks of sensors and actuators with the properties of fault tolerance when single failures occurred.

At the system level, we will develop diagnostic software for many types of spacecraft failures to ensure fault tolerance using systemic connections. Also at the system level, we will solve the problem of spacecraft orientation and stabilization system restore functionality in catastrophic failure in one of the blocks – BRW or CAVG through the use of auxiliary control devices (magnetometers, jet engines), and measuring devices (star trackers, GPS, etc.)

Failure modes were grouped into classes defined by the change of one of the system mathematical model parameters because of the failures parameterization. Set of failures classes contains 46 elements that are direct diagnostic features of failure.

Motion model of system in the presence of it failures obtained based on the equations system describing the dynamics of the stabilization and orientation system of the spacecraft and set of straight diagnostic features. In this case, is made the transition to a mathematical description in

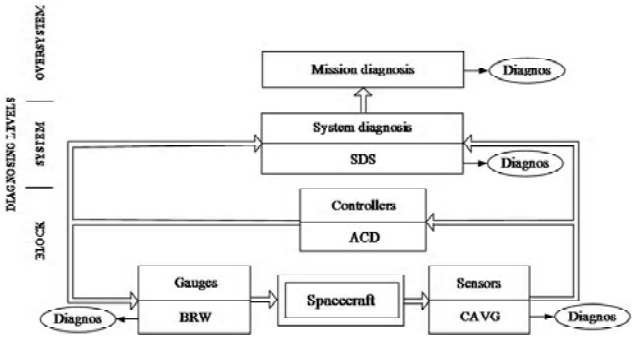


Fig. 5. Functional scheme of the hierarchy of diagnostic software

recurrence-matrix form of the system motion equations linearized at the operating point.

Analysis of the structural and signaling systems diagnosable relative to this set of failures classes showed that the complete diagnosable system is provided on transitive operating modes of RW and spacecraft in general and to differ PA and RW failures is necessary to introduce additional control point between them. Spacecraft is the object with a known input. Diagnosing its functional state based on the difference signal of measurement vector and the reference model output. Since all state variables of spacecraft (angular speed velocity of the spacecraft and RW) are directly measurable, as a function of diagnostic models that connect direct and indirect failure indications, selected models species

$$\begin{aligned}\Delta x(k+1) &= A\Delta x(k) + \Delta A(\Delta\gamma_q)\tilde{x}(k) + \Delta B(\Delta\gamma_q)u(k) + \\ &+\Delta F(\Delta\gamma_q)f(k) + \tilde{B}u_0(\Delta\gamma_q); \\ \Delta y(k) &= C\Delta\, x(k) + \Delta C(\Delta\gamma_q)\tilde{x}(k) + \tilde{C}u_0(\Delta\gamma_q).\end{aligned} \quad (2)$$

where $\Delta A(\Delta\gamma_q)$, $\Delta B(\Delta\gamma_q)$, $\Delta F(\Delta\gamma_q)$, $\Delta C(\Delta\gamma_q)$, are deviations respectively state matrices of control, disturbance, and output due to the presence of a direct indication of failure $\Delta\gamma_q$ in system; $\tilde{B}u_0(\Delta\gamma_q)$, $\tilde{C}u_0(\Delta\gamma_q)$ are matrix characterizing the influence of drift in the elements of the control object on its state variables and output signals.

Based on the obtained models developed procedures for processing indirect diagnostic fault features for block and system levels. They allow to consistently remove the uncertainties associated with the moment of a failure, its location, class, and the type. It was taken into account following. When establishing the facts failure, finding its place and the class definition is used dichotomous tree that represents the production knowledge base of diagnosing process. The nodes of this tree are used predicate constructions of double type

$$z = S_2\left\{f\left[\Delta y(k)\right] - \delta\right\} = \begin{cases}1, & \text{if } f[\cdot] \ge \delta; \\ 0, & \text{if } f[\cdot] < \delta;\ \forall k \in T,\end{cases} \quad (3)$$

where δ is threshold; $f\left[\Delta y_i(k)\right]$ is nonlinear function of measurement vector $\Delta y(k)$.

So, as the argument of double predicate uses discrete deviation of the output signals measurement of the system, placed in the vector $\Delta y(k)$, functional relationship of these measurements and the threshold value that determines function allowable change $f[\cdot]$. Nonlinear function $f[\cdot]$ formed via diagnostic models, connecting specific for each of the main diagnosing problem the direct feature $\Delta\lambda_i$ with indirect – the calculation results $\Delta y(k)$.

To provide diagnosability of the system with a depth to the RW functional element and differences between failures in spacecraft and AVS, for all operating modes introduce additional diagnostic features formed on the basis of the hypothesis of quasi-stationary direct feature of failure in the diagnosis interval:

$$\begin{aligned}&\Delta\hat{\gamma}_i(k) = \text{const} \cap \Delta\hat{\gamma}_j(k) = \text{var, if } \left(\Delta\hat{\gamma}_i \ne 0\right) \cap \left(\Delta\hat{\gamma}_j = 0\right); \\ &z_{Di} = s_2\left(\left|I_{CTi}\right| < \Delta I_{CTi}\right),\ I_{CTi} = \frac{d(\hat{\gamma}_i)}{dt}; \\ &i, j = \overline{1, N_\gamma}, i \ne j.\end{aligned} \quad (4)$$

Fig. 6, 7 and 8 are fragments of search algorithms for place and the class definition of failure.

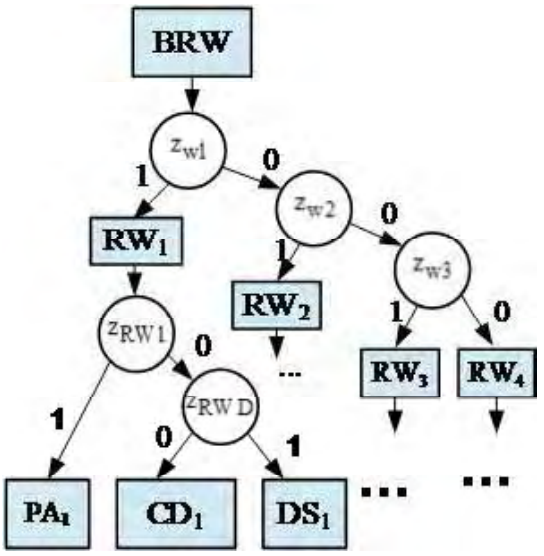


Fig. 6. Dichotomous tree search for a place of failure in the BRW

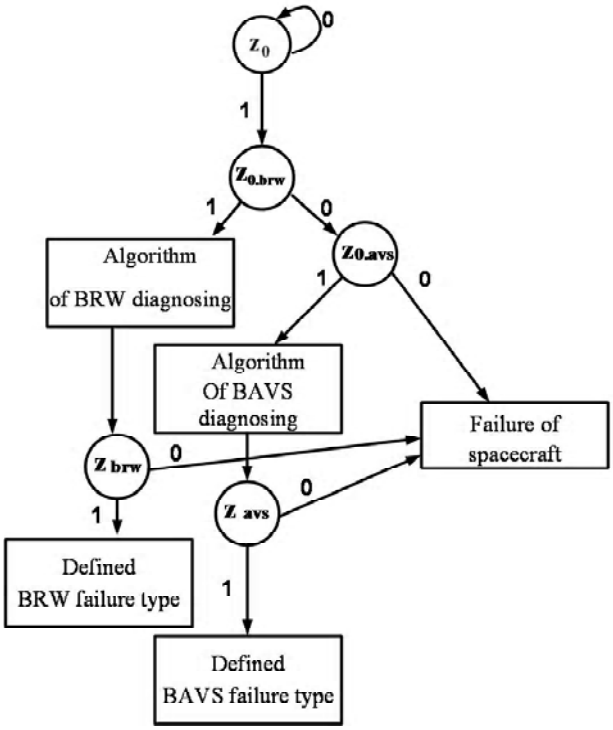


Fig. 7. Dichotomous tree search for a place of failure in the spacecraft orientation and stabilization system

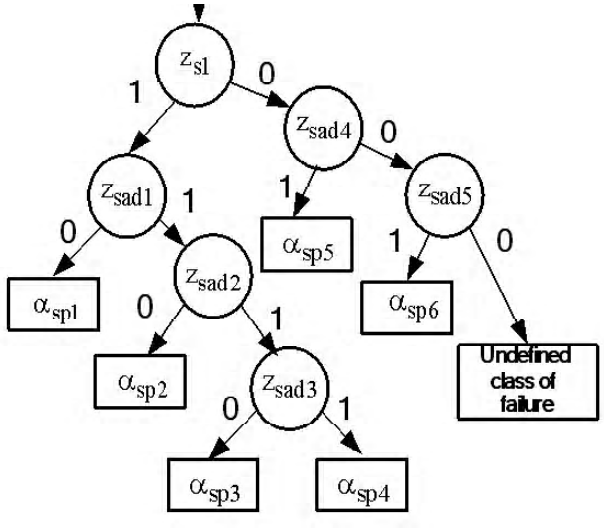


Fig. 8. Fragment of the dichotomous branch of the failure class definition in the spacecraft

To determine the failure mode are used functional dependencies, linking a particular failure type with direct diagnostic features. Final estimate of the parameter deviations value from the nominal system performed at the last stage of diagnosis procedure as a result of the evaluation characteristic values choice, calculated on the steps of determining the place and class of failure characterizing the detected type of failure

$$\Delta\hat{\gamma}_i(k) = \begin{pmatrix} CA_{\gamma_i} C^{-1}\tilde{y}(k-1) + C_{u0_{\gamma i}} + \\ +CB_{\gamma_i} u(k-1) + CF_{\gamma_i} f(k-1) + \\ +CB_{u0_{\gamma i}} + C_{\gamma i} C^{-1}\tilde{y}(k) \end{pmatrix}^{-1} \Delta y(k). \quad (5)$$

RECOVERY SYSTEM DESIGN

Developed diagnostic software is a foundation for the development of tools for the spacecraft orientation and stabilization system functional state automatic recovery in emergencies.

Consider the recovery of functional state of spacecraft orientation and stabilization system on block level. The starting point for this task is complete diagnosis of the function block. Based on this information, depending on the block construction, the principle of its action, the failures set are selected recovery tools – signal and parametric adjustment or possible hardware and control algorithms reconfiguration of the block.

Selection of recovery tools is produced by bases there control system functioning history, the current type of failure, the excess resources remaining and future challenges of the space mission. In this case, the recovery procedure related to the signal and parametric tuning often have the highest priorities, and procedures relating to the reconfiguration of the hardware and control algorithms are called only after all the resources associated with the adjustment have been exhausted.

After restoring the damaged unit is produced its diagnosis of the functional state. As a result:

– fully restored functionality of the block and it can be used in the orientation and stabilization system of spacecraft as intended;

– working capacity of the unit is not restored and detected another type of failure, and the previously described cycle is repeated until complete recovery the functional properties of the emergency unit.

After a complete recovery, the emergency unit on the system level signal is transmitted and produced parry of system failure effects in the blocks for the subsequent implementation of the flight mission.

However, such situation is possible when failures on block level can produce such abnormal situation, which is not possible to parry at this level because of the lack of excess capacity. In this case, the recovery procedure of the functional state of spacecraft occurs at the system level, using the connection between the block and the additional devices that are not intended directly for the tasks of orientation and stabilization of the spacecraft, but allowing performing tasks for space missions.

At the system level are also solved problem of flexible parry abnormal system situations associated with failures in the spacecraft construction or significant perturbing effects.

After the spacecraft orientation and stabilization system recovery are also produced diagnosing of the restored system. The full diagnosis allows make a decision to continue parry the effects of failures, or about the further implementation of the mission, or about the need parry on oversystem level (Fig. 9).

EXPERIMENT

Debugging algorithms of diagnosing functional state and recovery of the orientation and stabilization system of the spacecraft performed on a specialized hardware-software complex (HSC), presented in Fig. 10 and Fig. 11.

Functional structure of HSC allows you to enter into the system failure modes from the considered set, and to investigate the behavior of the system in the nominal and emergency modes.

HSC consists of the following elements: a platform with minimally redundant BRW, established by pyramid scheme and CAVG, fixed in gimbals, automatic control device and the PC with specialized software.

The complex software includes low-level software that provides execution in the control microcontroller module and the peripheral microcontroller module developed procedures control of and provide the active fault tolerance. On the upper level, there is software for performing the

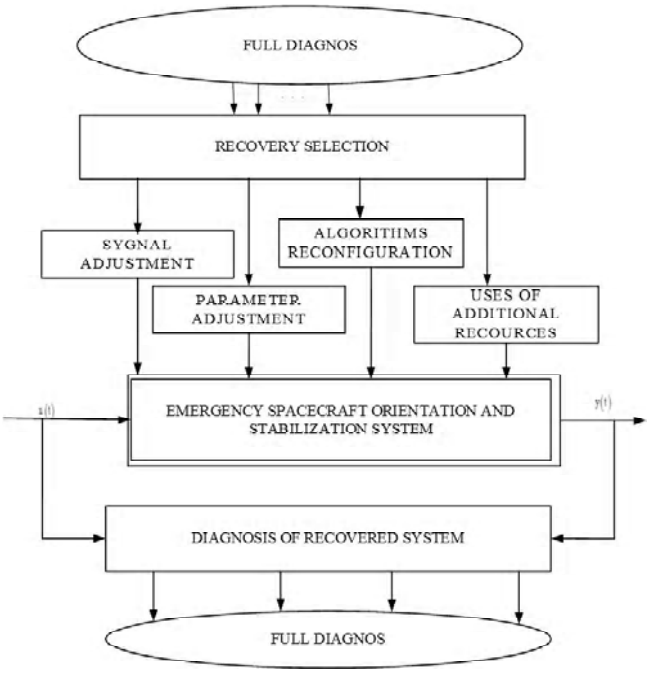


Fig. 9. Parrying tools of orientation and stabilization system of the spacecraft failures at the system level

following functions: information exchange with automatic control device, setting modes of the system, visualization and storage of the experimental results.

In order to simulate effect of the considered set of failure modes on the system using developed HSC, uses software simulators. They are used because of the fact that they have essential advantages such as the possibility of rapid

Fig. 10. Exterior view of HSC for research models and methods of maintenance the active fault-tolerance spacecraft attitude control and stabilization

modification, imitation of wide variety operating conditions, ease of the results interpretation, small power consumption, a good adaptability. At this distortion of measured and control signals of the system performed so that the response to it was analogous reaction to input types of failure.

Fig. 12 shows the functional diagram of RW with software simulators of failure modes.

Fig. 13 and 14 shows the results of the spacecraft stabilization control system simulation without active fault tolerance (dotted line) and with an activated diagnosis and recovery system (solid line) at the mode of the angular reorientation of the platform along the OZ axis by predetermined path.

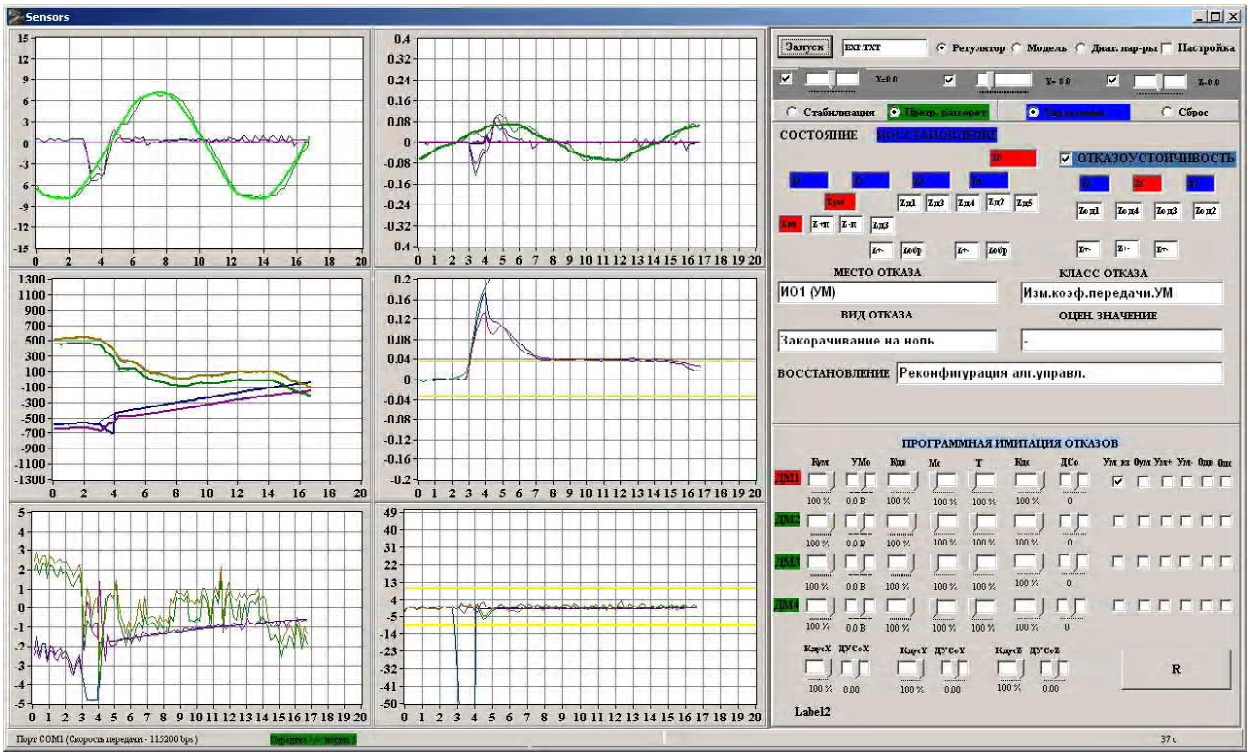


Fig. 11. The main window of the program results visualization and setting a system operation for spacecraft attitude control and stabilization

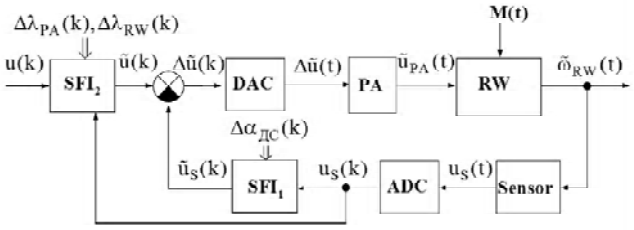


Fig. 12. The functional diagram of RW with software simulators of failure modes

Similarly, research has been done in the modes of attitude control and stabilization object functioning on the whole parrying failure modes set. All types of failure were observed by diagnostics subsystem, found their place, set the class and defined failure mode.

Average failure detection time was 0,07 s with the average total time of diagnosis – 0,32 s, and the average time failure parrying – 1,35 s. Thus, the average time spent on diagnosis and parry of spacecraft stabilization control system failures was 1,74 seconds. This time less than time t_{pp} = 9 s for transient system in normal mode in 5,17 times

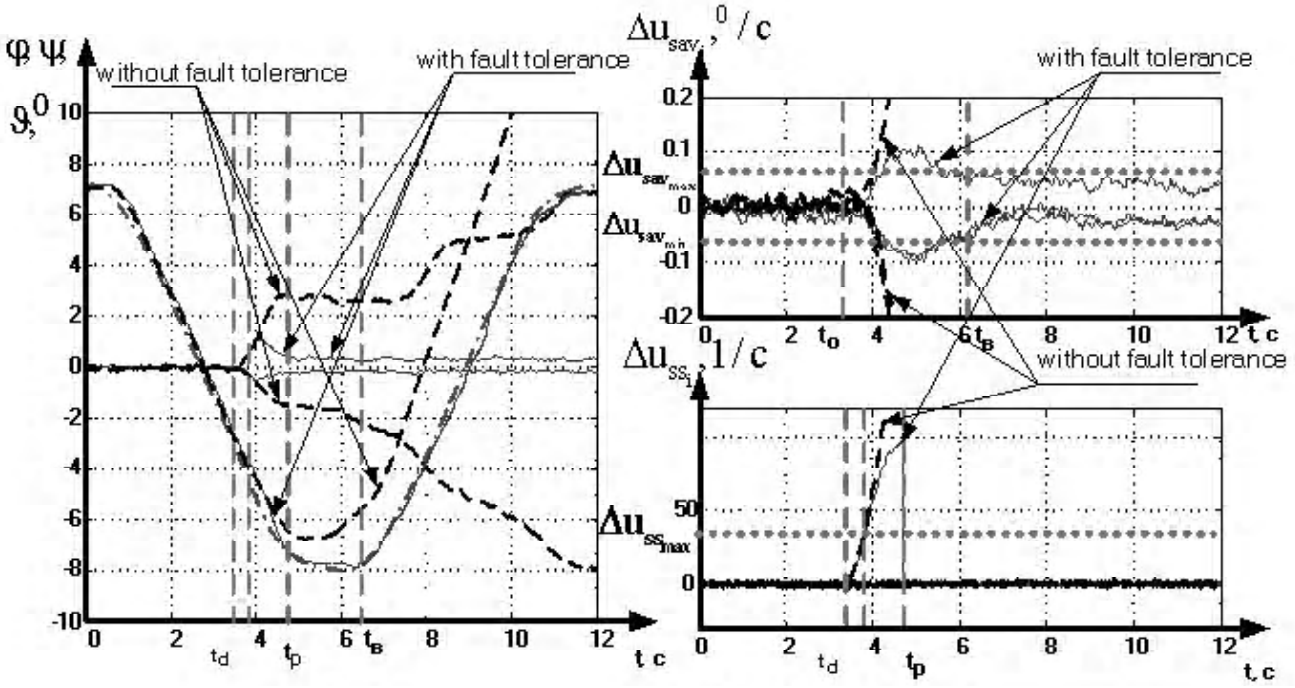


Fig. 13. The results of the spacecraft stabilization control system at break RW with a fourth serial number

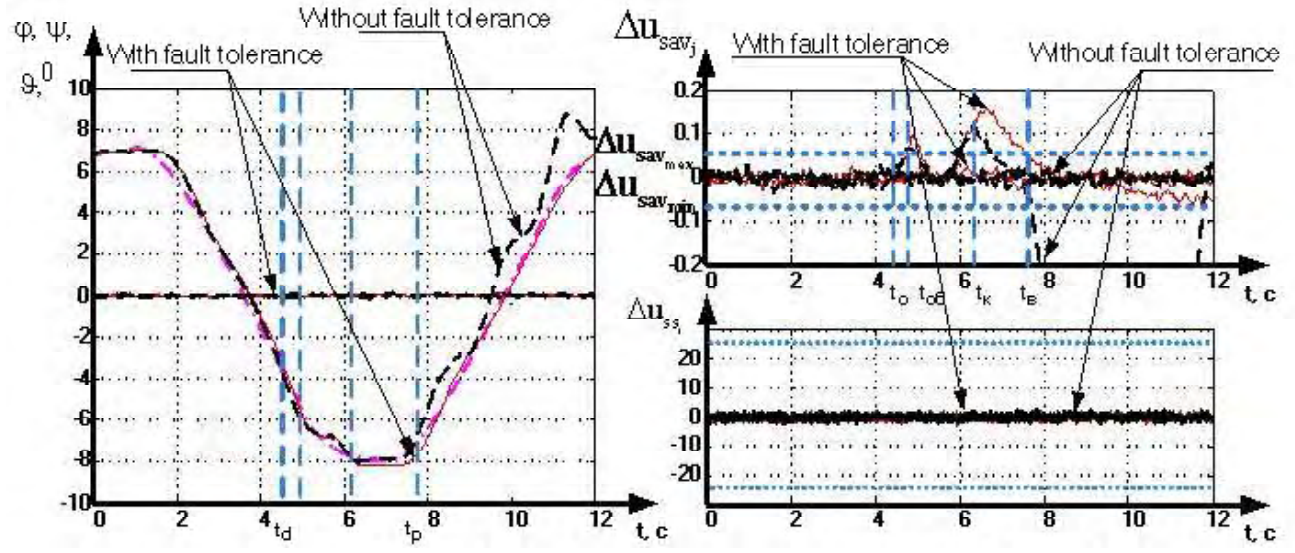


Fig. 14. The results of the spacecraft stabilization control system with a decrease in gain Kz

CONCLUSIONS

Thus in this paper identified the main trends in modern space technology that is in reducing the weight and size, power and cost characteristics, function of space vehicles, as well as increasing the time of their active life. There was formed the structure of the orientation and stabilization system for ensuring its active fault tolerance feature, defined a set of standard and parameterized types of failures of the system functional elements, characterized by changes in the properties of the conversion elements. There was received diagnostic functional models for solving problems of detection, search the place and establish a class of failure, which provide unique analytical relation between direct and indirect diagnostic features for each diagnosing problem, take into account the dynamic features of the object, as well as systemic linkages between its elements. There was developed diagnostic logic models for detection, finding a place and establish a class of failure in the system of automatic orientation and stabilization, the expressions for the calculation of direct and indirect diagnostic features, and formed their processing rules that al-low to solve problems of analytically develop procedures for deep diagnosing. Further developed models and methods of systematic approach to fault tolerance in the direction of the effective use of the signal, parametric and structural redundancies and selection of parrying tools when there are multiple connections between the means of parrying and failure modes. Experimental researches of the sample model of the automatic attitude control and stabilization the properties of active fault-tolerance in emergency modes of operation which have shown operability of the developed models and methods of deep diagnosing and failures flexible parrying and fundamental possibility maintain operability of the object in the event of failure modes in it from the finished set.

SPISOK LITERATURY

1. Проектирование и экспериментальная отработка систем управления объектов ракетно-космической техники. Т. 2. Проектирование систем управления космических аппаратов и модулей орбитальных станций [Текст]: учебник в 3 т. / Ю. С. Алексеев, Е. В. Белоус, Г. В. Беляев и др. / под общей ред. Ю. С. Алексеева, Ю. М. Златкина, В. С. Кривцова, А. С. Кулика, В. И. Чумаченко. – Х. : Нац. аэрокосм. ун-т им. Н. Е. Жуковского «Харьк. авиац. ин-т», НПП Хартрон-Аркос, 2012. – 680 с.
2. *Кулик, А. С.* Алгоритмическое обеспечение модулей диагностирования и восстановления работоспособности спутниковой системы ориентации и стабилизации [Текст] / А. С. Кулик, О. А. Лученко, С. Н. Фирсов // Радіоелектроніка, інформатика, управління. – 2012. – №1 (26). – С. 112–122.
3. *Постников, В. Н.* Критерии оценки свойств избыточных маховичных систем [Текст] / В. Н. Постников, А. Н. Таран, С. Н. Фирсов // Радіоелектронні та комп'ютерні системи. – 2011. – №4 (52). – С. 82–87.
4. *Кулик, А. С.* Концепция обеспечения живучести спутниковых систем управления ориентацией и стабилизацией / А. С. Кулик, О. А. Лученко, С. Н. Фирсов // Радіоелектроніка, інформатика, управління. – 2011. – № 2 (25). – С. 41–47.
5. *Кулик, А. С.* Эффективность избыточных систем стабилизации и ориентации космических аппаратов с двигателями-маховиками [Текст] / А. С. Кулик, А. М. Суббота, О. В. Резникова // Авиационно-космическая техника и технология. – 2008. – № 3 (50). – С. 18–25.
6. *Фирсов, С. Н.* Обеспечение функциональной устойчивости измерителей параметров движения спутниковых систем стабилизации и ориентации / С. Н. Фирсов // Радіоелектроніка, інформатика, управління. – 2013. – №1 (28). – С. 144–150.
7. *Кулик, А. С.* Восстановление измерений навигационной системы в режиме реального времени [Текст] / А. С. Кулик, С. Н. Фирсов, До Куок Туан, О. Ю. Златкин // Радіоелектронні і комп'ютерні системи. – 2008. – Вып. 5 (59). – С. 28–33.
8. *Фирсов, С. Н.* Обеспечение глубокого диагностирования блока акселерометров при плоскостном движении летательного аппарата [Текст] / С. Н. Фирсов, Туан Куок До // Радіоелектронні і комп'ютерні системи. – 2009. – Вып. 3 (60). – С. 33–38.

Стаття надійшла до редакції 16.09.2013.

Фирсов С. Н.[1], Резникова О. В.[2]

[1]Канд. техн. наук, доцент, Национальный аэрокосмический университет им. Н. Е. Жуковского «ХАИ», Украина

[2]Асистент, Национальный аэрокосмический университет им. Н. Е. Жуковского «ХАИ», Украина

ОБЕСПЕЧЕНИЕ ОТКАЗОУСТОЙЧИВОСТИ СИСТЕМЫ ОРИЕНТАЦИИ И СТАБИЛИЗАЦИИ КОСМИЧЕСКИХ АППАРАТОВ

Сформирована структура системы ориентации и стабилизации, обеспечивающая ее активную отказоустойчивость. Параметризированы виды отказов функциональных элементов системы, характеризующиеся изменением свойств преобразовательных элементов. Получены диагностические функциональные модели для решения задач обнаружения, поиска места и определения класса отказа системы. Разработаны модели и методы системного подхода к отказоустойчивости с целью эффективного использования сигнальной, параметрической и структурной избыточностей и выбраны инструменты парирования. Проведены экспериментальные исследования макетного образца автоматической системы ориентации и стабилизации для исследования свойств активной отказоустойчивости в аварийных режимах работы.

Ключевые слова: система стабилизации и ориентации космических аппаратов, отказоустойчивость, двигатель маховик, датчик угловой скорости.