

МАТЕМАТИКА

УДК 519.72

doi: 10.31429/vestnik-15-2-12-18

РАЗРАБОТКА ЧИСЛЕННОГО АЛГОРИТМА ГЕНЕРАЦИИ ПОРОЖДАЮЩИХ СИСТЕМ И ПОИСКА ОПТИМАЛЬНЫХ ИЗ НИХ

Синельникова Т. И.

DEVELOPMENT OF NUMERICAL ALGORITHM FOR GENERATING GENERATIVE SYSTEMS AND FOR THE OPTIMAL OF THEM SEARCHING

T. I. Sinelnikova

Kuban State University, Krasnodar, Russian Federation
e-mail: sinelnikova.email@mail.ru

Abstract. This article overviews models of the epistemological level of generative systems and their properties. The necessity to apply given systems and the theoretical basis of their formation are stated. A mathematical formalism of systemological methodology is applied for systems problem solving. The author gives a detailed description of the developed numerical algorithm of the generation and search of optimal generative systems. The author's algorithm provides execution of requirements shown to meaningful submasks and corresponding generative systems in systemology. It also fulfills the requirement for optimal generative systems. This article features the results of the algorithm testing on the data system with probability distribution functions. In work the comparative analysis of generative systems with various parameters is carried out, as well as comparison of the results got by means of the author's algorithm with the results given by other authors are presented. On the basis of obtained results was drawn a conclusion that efficiency of the offered algorithm.

Keywords: systemology, generative systems, numerical methods for systems problem solving.

Системологическая методология базируется на классификации систем и методов в иерархии эпистемологических уровней [1]. Моделирование систем включает переход к общим системам-представителям путем последовательного абстрагирования исследуемой конкретной задачи. Решение системных задач и основополагающих задач системологии осуществляется на третьем уровне иерархии — уровне структурированных систем, который включает подъем по нижестоящим уровням и характеризует отношения между математическими моделями систем предыдущих уровней [2]. Однако на предшествующем ему эпистемологическом уровне отсутствует численный алгоритм генерации и поиска оптимальных порождающих систем на основе моделей систем данных первого уровня.

Для реализации математических методов решения системных задач второго и третьего уровня разработан авторский численный алгоритм генерации порождающих систем и поиска оптимальных из них. Начальными условиями служат параметры наибольшей допу-

стимой маски и требования, предъявляемые к данному классу систем.

В основе порождающих систем лежат системы с поведением. Термин “поведение” характеризует общее инвариантное на параметрическом множестве W ограничение на переменные множества V общей представляющей системы I и абстрактные переменные, которые определяются на параметрическом множестве правилами сдвига. Если элементы полностью упорядоченного параметрического множества T представляют собой последовательные целые положительные числа, то правило сдвига задается уравнением: $r_j(t) = t + \rho_j$, $t \in T$, где $j \in N_m$ ($N_m = \overline{1, m}$, m — число параметров), ρ_j — константа $\rho_j \in Z$ (Z — множество целых чисел). Результатом применения набора R таких правил сдвига является множество выборочных переменных, характеризующихся маской M , то есть множеством, которое определяется отношением $M \subseteq V \times R$, таким что $(\forall (v_i, r_j) \in M) \rightarrow s_{k,w}$, где $v_i \in V$, $i \in N_n$ (n — число базовых переменных, входящих в

систему), $r_j \in R$, $s_{k,w}$ — состояние выборочной переменной s_k при значении параметра w ($w \in W$). Множество выборочных переменных имеет вид: $S = \{s_1, s_2, \dots, s_k\}$, где $k \in N_{|M|}$, $|M|$ — мощность множества M . Полное множество состояний выборочных переменных представляет собой декартово произведение $C = S_1 \times S_2 \times \dots \times S_k$. Функция $f_B : C \rightarrow \{0,1\}$, определяющая реально встречающиеся состояния c , называется функцией поведения. Она формируется посредством системы данных D и маски, которая, в свою очередь, определяется через переменные и параметры системы I .

Порождающие системы с поведением [1]

$$F_{GB} = (I, M_G, f_{GB}),$$

где $M_G = (M, M_g, M_{\bar{g}})$ — маска порождения: маска и ее разбиение на порождаемую и порождающую подмаски (для направленных систем M_G включает M_e и $M_{\bar{e}}$ для множества входных E и выходных \bar{E} переменных); f_{GB} — выражает способ представления состояния порождаемых переменных ($g \in G$) по состоянию порождающих переменных ($\bar{g} \in \bar{G}$).

Все содержательные подмаски наибольшей допустимой маски M образуют ограниченное множество систем с поведением. Подмаска считается содержательной, если она удовлетворяет следующим требованиям [1]:

1) содержит хотя бы по одному элементу из каждой подмаски M_i ($M_i = \{(\alpha, \beta) | (\alpha, \beta) \in M, \alpha = v_i\}$);

2) содержит хотя бы один элемент с правилом сдвига $t + \max(\rho)$ при перемещении маски по матрице данных в направлении возрастания параметра или $t + \min(\rho)$ при перемещении маски в направлении убывания параметра.

Таким образом, число возможных подмасок [1]

$$N = (2^{\Delta M} - 1)^n - (2^{\Delta M - 1} - 1)^n,$$

где n — число базовых переменных; $\Delta M = 1 + \max(\rho) - \min(\rho)$ — глубина маски M .

Применение порождающих систем на следующем эпистемологическом уровне требует, чтобы они наилучшим образом представляли набор эмпирических данных, следовательно, выбор должен быть осуществлен в пользу систем, которые имеют наименьшую порождающую нечеткость. Для вероятностных систем

порождающая нечеткость H представляет собой шенноновскую энтропию [3] с множествами выходов C, G, \bar{G}, E и распределениями вероятностей, представленными соответствующими функциями поведения; в случае исследования систем с функциями распределения возможностей порождающая нечеткость представляет собой U -нечеткость [1]. К порождающим системам также предъявляются требования согласованности, детерминированности, простоты и соответствия [1]. Таким образом, для поиска оптимальных из них требуется провести исследование всех содержательных подмасок заданной наибольшей допустимой маски. Выбор наиболее информативных свойств системы, также входит в область задач методов исследуемого уровня. Выявление значимых переменных на основе моделирования численными методами порождающих систем дает возможность экономического представления данных [4].

Создан численный алгоритм, позволяющий произвести генерацию порождающих систем, осуществить выбор их оптимальных вариантов, удовлетворяющих указанным выше требованиям, и включает следующую последовательность этапов:

- 1) задать направление маски;
- 2) определить $V, \Delta M, |M|$ ($|M|$ — мощность M);
- 3) произвести разбиение E ;
- 4) если $|E| > 0$ ($|E|$ — мощность E), то задать наличие информации о переменных системы,

$$j = \begin{cases} 0, & \text{если информация есть,} \\ 1, & \text{если информации нет;} \end{cases}$$

- 5) генерация множеств масок L ;
- 6) задать $i = 0, m = |M|$;
- 7) задать $k = 0, l = |L_i|$, где $L_i \subset L$ ($|L_i|$ — мощность L_i);
- 8) произвести разбиение M_k на $M_g, M_{\bar{g}}, M_e$, где $M_k \in L_i$ и

$$M = \begin{cases} \bigcup_{n \in N_{|V|}} (\{s_{(n-1)\Delta M + 1}, \dots, s_{(n-1)\Delta M + \Delta M}\}), & \text{если } t \text{ возрастает,} \\ \bigcup_{n \in N_{|V|}} (\{s_{(n-1)\Delta M + \Delta M}, \dots, s_{(n-1)\Delta M + 1}\}), & \text{если } t \text{ убывает;} \end{cases}$$

9) вычислить нечеткость:

$$H_k = \begin{cases} H(G|\bar{G}) = H(C) - H(\bar{G}) = \\ \quad = - \sum_{c \in C} f(c) \log_2 f(c) + \\ \quad \quad + \sum_{\bar{g}} f(\bar{g}) \log_2 f(\bar{g}), \\ \quad \quad \quad \text{если } |E| = 0; \\ H(G|E \times \bar{G}) = H(C) - H(E \times \bar{G}), \\ \quad \quad \quad \text{если } j = 0, \\ H(G|E \times \bar{G}) = \\ \quad = \frac{1}{|E|} \left(\sum_{e \in E} H(\bar{E}|e) - \sum_{e \in E} H(\bar{G}|e) \right), \\ \quad \quad \quad \text{если } j = 1; \end{cases}$$

10) если $k = 0$ или $H_k < H_{\min}$, то $H_k \rightarrow H_{\min}$, $M_k \rightarrow M_{\min}$;

11) если $k < l$, то $k + 1 \rightarrow k$, перейти на шаг 8;

12) $(H_{\min}, M_{\min}) \rightarrow X_i$, записать пару X_i как оптимальную пару уровня сложности i ;

13) если $i < (m - 2)$, то $i + 1 \rightarrow i$, перейти на шаг 7, иначе перейти на шаг 14;

14) конец.

Для возможностных порождающих систем на девятом шаге вычисляется U -нечеткость.

В вышеизложенный алгоритм входит авторский алгоритм генерации масок (шаг 5), включающий следующие этапы:

1) задать $m = \Delta M |V|$ ($|V|$ — мощность V), $M_0 = \{s_1, s_2, \dots, s_n\}$, $n = \overline{1, m}$, $M_0 \rightarrow L_0$;

2) задать $i = 0$;

3) задать $q = 0$, $l = |L_i|$ ($|L_i|$ — мощность L_i);

4) задать $k = 1$, $r = 0$;

5) если $s_k \in M_q$, где $M_q \in L_i$, то $M_q \setminus \{s_k\} \rightarrow M_r$;

6) если $M_r \neq \emptyset$, $M_r \notin \bigcup_{y \in Y} L_y$, где $Y = \{y \mid y \in Z, 0 \leq y \leq i\}$, и M_r удовлетворяет требованиям содержательной маски, то записать маску M_r в L_{i+1} , $r + 1 \rightarrow r$;

7) если $k < m$, то $k + 1 \rightarrow k$, перейти на шаг 5;

8) если $q < (l - 1)$, то $q + 1 \rightarrow q$, перейти на шаг 4;

9) записать L_{i+1} , если $i < (m - 2)$, то $i + 1 \rightarrow i$ перейти на шаг 3, иначе на шаг 10;

10) конец.

Разработанный численный алгоритм позволил применить методы эпистемологического уровня порождающих систем и осуществить их программную реализацию [5]. Алгоритм апробирован на эмпирических данных из исследований, проводимых в работах [1, 6]. Исследовалась нейтральная система данных (табл. 1), описывающая поведение двух птиц за 90 секундный период видеофиксации, который был разделен на двухсекундные интервалы, в течение которых проводилась этологическая оценка поведения птиц и соотнесение его с четырьмя выделенными состояниями (от 1 до 4) [6].

В результате работы алгоритма для наибольшей допустимой маски (НДМ) с вероятностной мерой нечеткости и параметрами: $\Delta M = 2$, направление движение маски (1) — в сторону возрастания параметра t , получено 8 систем: на 4-м уровне сложности сформирована одна система, маска оптимальной системы №1, порождающая нечеткость $H = 1,087$; на 3-м уровне сгенерировано 4 системы, маска оптимальной системы №3, порождающая нечеткость $H = 1,897$; на 2-м уровне сгенерировано 3 системы, маска оптимальной системы №1, порождающая нечеткость $H = 3,396$. При направлении движения маски (2) — в сторону убывания t , с теми же параметрами НДМ, получено 8 систем: на 4-м уровне сложности сформирована одна система, маска оптимальной системы №1, порождающая нечеткость $H = 1,2$; на 3-м уровне сгенерировано 4 системы, маска оптимальной системы №3, порождающая нечеткость $H = 1,912$; на 2-м уровне сгенерировано 3 системы, маска оптимальной системы №2, порождающая нечеткость $H = 3,429$.

Результаты, полученные посредством численного алгоритма, для матрицы данных (табл. 1) и НДМ с параметрами: $\Delta M = 3$, направление (1), получено 40 систем: на 6-м уровне сложности сформирована одна система, маска оптимальной системы №1, порождающая нечеткость $H = 0,413$; на 5-м уровне сгенерировано 6 систем, маска оптимальной системы №1, порождающая нечеткость $H = 0,552$; на 4-м уровне сгенерировано 14 систем, маска оптимальной системы №5, порождающая нечеткость $H = 1,067$; на 3-м уровне сгенерировано 14 систем, маска оптимальной системы №6, порождающая нечеткость $H = 1,882$; на 2-м уровне сгенерировано 5 систем, маска оптимальной системы №1, по-

Таблица 1. Матрица данных этологической системы

t, c	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32
v_1	1	1	0	3	3	3	3	3	4	3	3	0	2	1	1	1
v_2	4	3	4	3	3	3	3	3	4	4	3	3	4	2	1	1
t, c	34	36	38	40	42	44	46	48	50	52	54	56	58	60	62	64
v_1	1	4	4	4	4	4	0	2	2	4	4	4	4	4	4	4
v_2	1	4	4	4	4	3	1	4	2	2	4	4	4	4	4	4
t, c	66	68	70	72	74	76	78	80	82	84	86	88	90			
v_1	2	2	2	2	2	0	2	1	0	2	1	1	4			
v_2	4	4	1	1	1	3	3	1	3	3	3	3	4			

рождающая нечеткость $H = 3,379$. Для НДМ с параметрами: $\Delta M = 3$, направление (2), получено 40 систем: на 6-м уровне сложности сформирована одна система, маска оптимальной системы №1, порождающая нечеткость $H = 0,413$; на 5-м уровне сгенерировано 6 систем, маска оптимальной системы №6, порождающая нечеткость $H = 0,596$; на 4-м уровне сгенерировано 14 систем, маска оптимальной системы №4, порождающая нечеткость $H = 1,067$; на 3-м уровне сгенерировано 14 систем, маска оптимальной системы №13, порождающая нечеткость $H = 1,882$; на 2-м уровне сгенерировано 5 систем, маска оптимальной системы №4, порождающая нечеткость $H = 3,413$.

Для НДМ с параметрами: $\Delta M = 4$, направление (1), получено 176 систем: на 8-м уровне сложности сформирована одна система, маска оптимальной системы №1, порождающая нечеткость $H = 0,244$; на 7-м уровне сгенерировано 8 систем, маска оптимальной системы №6, порождающая нечеткость $H = 0,244$; на 6-м уровне сгенерировано 27 систем, маска оптимальной системы №11, порождающая нечеткость $H = 0,292$; на 5-м уровне сгенерировано 50 систем, маска оптимальной системы №13, порождающая нечеткость $H = 0,418$; на 4-м уровне сгенерировано 53 системы, маска оптимальной системы №43, порождающая нечеткость $H = 0,744$; на 3-м уровне сгенерировано 30 систем, маска оптимальной системы №10, порождающая нечеткость $H = 1,824$; на 2-м уровне сгенерировано 7 систем, маска оптимальной системы №1, порождающая нечеткость $H = 3,296$. Для НДМ с параметрами: $\Delta M = 4$, направление (2), получено 176 систем: на 8-м уровне сложности сформирована одна система, маска оптимальной системы №1, порождающая нечеткость $H = 0,244$; на 7-м уровне сгенерировано 8

систем, маска оптимальной системы №3, порождающая нечеткость $H = 0,244$; на 6-м уровне сгенерировано 27 систем, маска оптимальной системы №16, порождающая нечеткость $H = 0,292$; на 5-м уровне сгенерировано 50 систем, маска оптимальной системы №40, порождающая нечеткость $H = 0,477$; на 4-м уровне сгенерировано 53 системы, маска оптимальной системы №49, порождающая нечеткость $H = 0,837$; на 3-м уровне сгенерировано 30 систем, маска оптимальной системы №29, порождающая нечеткость $H = 1,871$; на 2-м уровне сгенерировано 7 систем, маска оптимальной системы №6, порождающая нечеткость $H = 3,396$.

Представленный численный алгоритм на каждом уровне сложности генерирует множество масок, на базе которых формирует порождающие системы, удовлетворяющие всем предъявляемым к ним в системологии требованиям, и осуществляет поиск их оптимальных вариантов. В обоих направлениях прохода маски на каждом уровне сложности выявлено только по одной оптимальной маске. График (рис. 1) представляет собой обобщение приведенных выше результатов вычислительного эксперимента по исследованию порождающих систем с различными параметрами и отражает изменения нечеткости порождающих систем в зависимости от уровня сложности маски. Сравнительный анализ изменения характеристик порождающих систем, полученных для различных направлений прохода маски по матрице данных, показал, что для масок с заданными параметрами направление (1) предпочтительнее, так как при нем порождающая нечеткость меньше. Следовательно, для исследуемой этологической системы легче предсказывать будущие состояния системы, чем определять прошлые. Если есть возможность выбрать один из по-

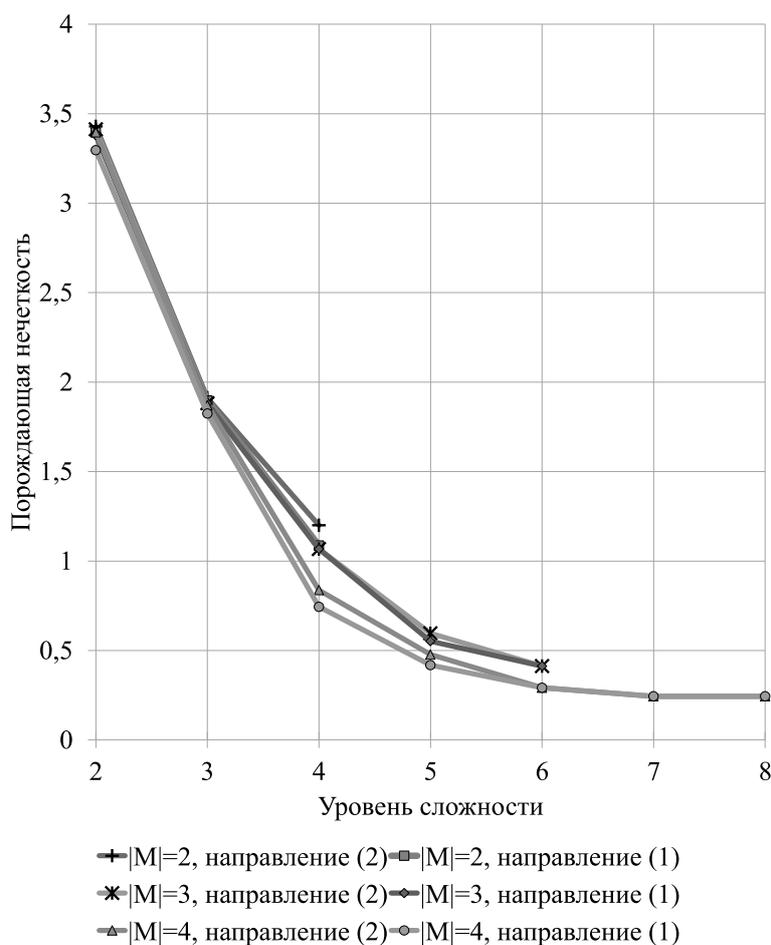


Рис. 1. Изменение характеристик порождающих систем

рядков следования, выбор стоит сделать в пользу первого.

Оценка полученных порождающих систем также показала, что с параметрами заданной наибольшей допустимой маски и $\Delta M = 4$ для исследования системы и порождения данных без потери информации может быть использована маска с меньшей сложностью (направление (1), маска №6, $|M| = 7$, $H(G|\bar{G}) = 0,244$; направление (2), маска №3, $|M| = 7$, $H(G|\bar{G}) = 0,244$).

На рис. 2, 3 представлены графики зависимостей, которые являются иллюстрацией сравнительного анализа результатов проведенного вычислительного эксперимента с известными результатами [1] и показывают высокую точность разработанного численного алгоритма. Так, сравнительный анализ показал, что отличие порождающей нечеткости составляет порядка $1 \cdot 10^{-2}$, наличие которого обусловлено изменением пространствен-

ного представления масок — расположением выборочных переменных и их нумерацией. При этом на выбор порождающих систем данное отличие не влияет; системы, полученные в результате работы алгоритма, полностью совпадают с результатами в работе [1], для представленных там случаев. Предложенный численный алгоритм является эффективным средством генерации порождающих систем и поиска оптимальных из них.

Полученные маски могут быть использованы для порождения состояний системы при заданном параметрическом множестве, что является необходимым при исследовании систем в условиях недостающей информации о состояниях их свойств. Применение аппарата порождающих систем позволяет определить оптимальное представление данных с точки зрения пространственных связей для дальнейшего исследования систем; кроме того, порождающие системы, основанные на струк-

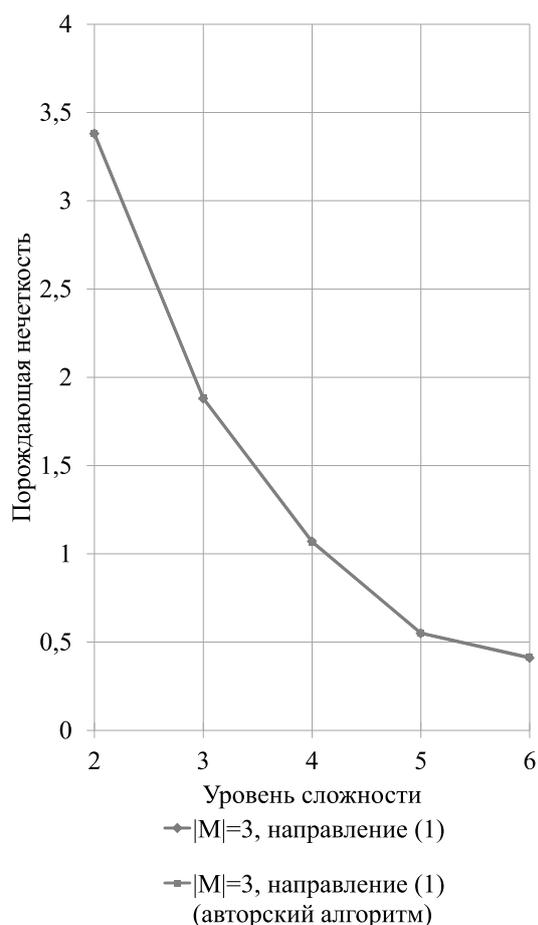


Рис. 2. Изменение характеристик порождающих систем, полученных посредством авторского алгоритма и представленных в работе [1] ($\Delta M = 3$)

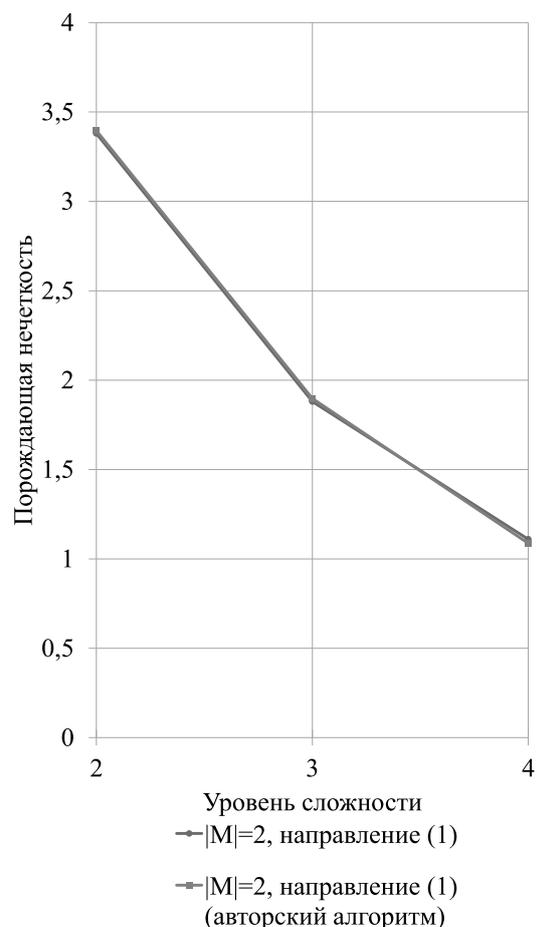


Рис. 3. Изменение характеристик порождающих систем, полученных посредством авторского алгоритма и представленных в работе [1] ($\Delta M = 2$)

туре связей между параметрами и переменными (в том числе входными и выходными), которую представляет маска, позволяют произвести выбор существенных свойств объекта. Анализ значимости свойств системы производится исходя из структуры связей между элементами оптимальных порождающих систем, сформированных из системы данных по маскам различного уровня сложности с учетом методологических отличий.

Из полученных результатов следует, что сложность маски прямо пропорциональна порождающей нечеткости системы, сформированной на ее основе. Связано это с тем, что в сложных масках используется большее количество информации. При этом выявлено, что с увеличением глубины маски: растет сложность вычислительной обработки выборочных переменных; стремительно возрастает

число содержательных подмасок; уменьшается опора на эмпирическую базу исследования (число неполных выборок равно $2(\Delta M - 1)$); возрастает число элементов, составляющих начальное условие, определение которого иногда бывает затруднено.

Таким образом, в представленной работе:

- разработан численный алгоритм генерации порождающих систем и поиска оптимальных из них для различных уровней сложности, созданный алгоритм включает набор параметров: параметры наибольшей допустимой маски; тип систем (нейтральные, направленные); требования предъявляемые к содержательным помаскам и др.;

- проведен численный эксперимент по исследованию порождающих систем с различными параметрами;

– подтверждено, что с увеличением сложности маски уменьшается порождающая нечеткость системы, при этом возрастает количество вычислений;

– выявлено, что для исследуемой в работе системы легче предсказывать будущие состояния системы, чем определять прошлые

– полученные инвариантные на параметрическом множестве ограничения могут быть использованы для порождения состояний переменных.

Литература

1. *Клир Дж.* Системология. Автоматизация решения системных задач. М.: Радио и связь, 1990. 535 с.
2. *Синельникова Т.И.* О целесообразности использования компьютерных структурированных систем в процессе выработки оптимальных управленческих решений // Экономика знаний: проблемы управления формированием и развитием: материалы VI Международной научно-практической конференции. Краснодар: Кубанский гос. ун-т., 2014. С. 317–325.
3. *Shannon C.E.* Communication Theory of Secrecy Systems // *Bell System Technical Journal*. 1949. Vol. 28. No 4. P. 656–715.
4. *Синельникова Т.И., Швецова Н.А.* Реализация методов системологии эпистемологического уровня порождающих систем // Научно-технический вестник Поволжья. 2016. №5. С. 209–211.
5. *Синельникова Т.И., Швецова Н.А.* Программа выбора оптимального набора существенных характеристик исследуемого сложного объекта на основе нечетких мер системологии // Федеральная служба по интеллектуальной собственности. Свидетельство о государственной регистрации программы для ЭВМ № 2016660456; зарегистрировано в Реестре программ для ЭВМ 15.09.2016 г.
6. *Timbergen N.* The Evolution of Behavior in Gulls // *Scientific American*. 1960. Iss. 6. Vol. 203. P. 118–130.

References

1. *Klir, G.* *Sistemology. Automation of system problems solving. Architecture of systems problem solving.* Radio i svyaz', Moscow, 1990. (In Russian).
2. *Sinelnikova, T.I.* The appropriateness of the use of computer structure systems in the development of optimal management decisions. *Proc. of rep. 6th int. conf. 'Economic knowledge: Problems of management of formation and development'*, 2014. pp. 317–325. (In Russian)
3. *Shannon, C.E.* Communication Theory of Secrecy Systems. *Bell System Technical Journal*. 1949, vol. 28, no 4, pp. 656–715.
4. *Sinelnikova, T.I., Shvetsova, N.A.* Realization of systemological methods of the epistemological level of generate systems. *Nauchno-tekhnicheskii vestnik Povolzh'ya* [Scientific and technical journal of the Volga region], 2016, no. 5, pp. 209–211. (In Russian)
5. *Sinelnikova, T.I., Shvetsova, N.A.* *Program to select optimal set of essential characteristics of composite objects on the basis of fuzzy measures.* The Russian Federal Service for Intellectual Property. Certificate of state registration of computer programs no.2016660456; Registered in the Register of Computer Programs 15.09.2016. (In Russian)
6. *Timbergen N.* The Evolution of Behavior in Gulls. *Scientific American*, 1960, vol. 203, iss. 6, pp. 118–130.