

РАЗРАБОТКА АЛГОРИТМОВ И ПРОГРАММ ДЛЯ ПОСТРОЕНИЯ АНАЛИТИЧЕСКОГО ОПИСАНИЯ ПРОЦЕССОВ НА ОСНОВЕ ЭВОЛЮЦИОННОГО ПОИСКА

Шилов А.К., Шилов Д.А., Селянкин В.В.

Таганрогский государственный радиотехнический университет
347928, Россия, г. Таганрог, Некрасовский 44
Телефон: (86344) 310885, E-mail: t_sda@mail.ru

В докладе предлагается один из методов автоматического построения аналитической модели процесса, использующий самоорганизацию. Алгоритмы, использующие самоорганизацию, широко применяются для решения задач структурной идентификации и прогнозирования сложных систем [1-5]. Данный метод может быть применен для аппроксимации кривых и поверхностей, что может использоваться при конструировании аналитического описания процессов и сжатии сигналов. В этом методе процесс генерации аналитической модели состоит из четырех этапов.

На первом этапе производится генерация сложных функций за счет использования базовых функций и функциональных конструкторов (верхняя часть рис. 1).

ПРИМЕР СИСТЕМЫ ПОСТРОЕНИЯ АНАЛИТИЧЕСКОГО ВИДА ФУНКЦИИ ОДНОЙ ПЕРЕМЕННОЙ

Базовые функции: $x+c$, $c-x$, $x*c$, c/x .

Описание конструкторов функций :

- 1) Новая сложная функция имеет вид - $Sf(\dots,c) = f(\dots)+c$, если последняя операция в функции родителе * или /.
- 2) Новая сложная функция имеет вид - $Sf(\dots,x) = f(\dots)-x$, если последняя операция в функции родителе * или /.
- 3) Новая сложная функция имеет вид - $Sf(\dots,c) = f(\dots)*c$, если последняя операция в функции родителе + или -.
- 4) Новая сложная функция имеет вид - $Sf(\dots,c) = c/f(\dots)$, если последняя операция в функции родителе + или -.

ДИАГРАММА КОНЕЧНОГО АВТОМАТА - РАСПОЗНАВАТЕЛЯ ЦЕПОЧЕК СИМВОЛОВ

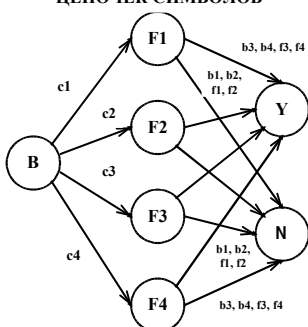
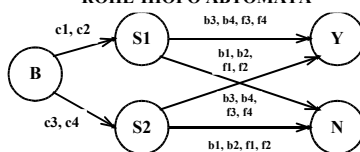


ДИАГРАММА ОПТИМИЗИРОВАННОГО КОНЕЧНОГО АВТОМАТА



В - начальное состояние; Y - конечное состояние соответствующее правильной цепочке; N - конечное состояние соответствующее неправильной цепочке; F1, F2, F3, F4 - состояния соответствующие выбору функциональных конструкторов; S1, S2 - эти же состояния после оптимизации.

Примеры правильных цепочек: c1 b3, c3 f2. Примеры неправильных цепочек: c2 b1, c4 f1.

Функц. констр.	Символ для КА
$f()+c$	c1
$f()-x$	c2
$f()*c$	c3
$c/f()$	c4

Функция	Символ для КА
$x+c$	b1
$c-x$	b2
$x*c$	b3
c/x	b4
$(\dots)+c$	f1
$(\dots)-x$	f2
$(\dots)*c$	f3
$c/(\dots)$	f4

Рис. 1

Задачей второго этапа является уменьшение перебора аналитических моделей. Для этой цели могут быть использованы конечные автоматы (нижняя часть рис. 1), формальные грамматики или таблицы с бинарными значениями (рис. 2). Во всех этих случаях процесс генерации аналитической модели может быть оптимизирован [7].

Кроме этого, можно выполнить уменьшение перебора моделей и с использованием алгоритма селекции (рис. 3). Алгоритмы селекции принадлежат более широкому классу эволюционных алгоритмов [1-3]. Алгоритмы селекции моделируют процесс искусственного отбора, который используется для выведения животных и растений со специфическими свойствами.

На третьем этапе осуществляется поиск параметров аналитической модели. Для этого используются различные способы нелинейной многомерной оптимизации, например генетические алгоритмы (рис. 4). Они моделируют процесс естественной биологической эволюции на уровне генов и хромосом. Причем одна хромосома включает в себя все параметры аналитической модели [6].

Последним этапом предлагаемого метода является процедура сопоставления некоторых аналитических моделей текущим порциям исходных данных. Для этого могут быть использованы алгоритмы кластеризации.

Алгоритм, построенный по предлагаемому методу, имеет точность около пяти процентов, когда он используется для функции имеющей пять постоянных и один переменный параметр (рис. 5).

Предлагаемый метод более гибок, чем его аналоги, так как он работает с аналитическим описанием почти любого вида. Кроме того, он может использоваться в таких приложениях, где другие алгоритмы дают плохие результаты.

ТАБЛИЦА ДЛЯ ОТСЕЧЕНИЯ НЕПОДХОДЯЩИХ СОЧЕТАНИЙ
ФУНКЦИОНАЛЬНЫХ КОНСТРУКТОРОВ И ФУНКЦИЙ-ПАРАМЕТРОВ

Конструкторы для сложных функций	Функции параметры							
	Базовые функции				Сложные функции			
	x+c	c-x	x*c	c/x	(...)+c	(...)-x	(...)*c	c/(...)
f()+c	0	0	1	1	0	0	1	1
f()-x	0	0	1	1	0	0	1	1
f()*c	1	1	0	0	1	1	0	0
c/f()	1	1	0	0	1	1	0	0

ТА ЖЕ ТАБЛИЦА ПОСЛЕ ОПТИМИЗАЦИИ

Конструкторы для сложных функций	Функции параметры			
	x+c, (...)+c	c-x, (...)-x	x*c, (...)*c	c/x, c/(...)
f()+c, f()-x	0	0	1	1
f()*c, c/f()	1	1	0	0

ФОРМАЛЬНАЯ КОНТЕКСТНО-СВОБОДНАЯ ГРАММАТИКА ДЛЯ ПОСТРОЕНИЯ АНАЛИТИЧЕСКОГО ВИДА ФУНКЦИИ ОДНОЙ ПЕРЕМЕННОЙ

Здесь "x", "c", "+", "-", "*", "/", "(", ")" - это терминалы;

<BF1>, <BF2>, <BF3>, <BF4> - это нетерминалы относящиеся к базовым и сложным функциям (функциям - параметрам);

<FC1>, <FC2>, <FC3>, <FC4> - это нетерминалы относящиеся к функциональным конструкторам.

<BF1> ::= x+c | <FC1>
 <BF2> ::= c-x | <FC2>
 <BF3> ::= x*c | <FC3>
 <BF4> ::= c/x | <FC4>

<FC1> ::= (<BF3>)+c | (<BF4>)+c
 <FC2> ::= (<BF3>)-x | (<BF4>)-x
 <FC3> ::= (<BF1>)*c | (<BF2>)*c
 <FC4> ::= c/(<BF1>) | c/(<BF2>)

Рис. 2

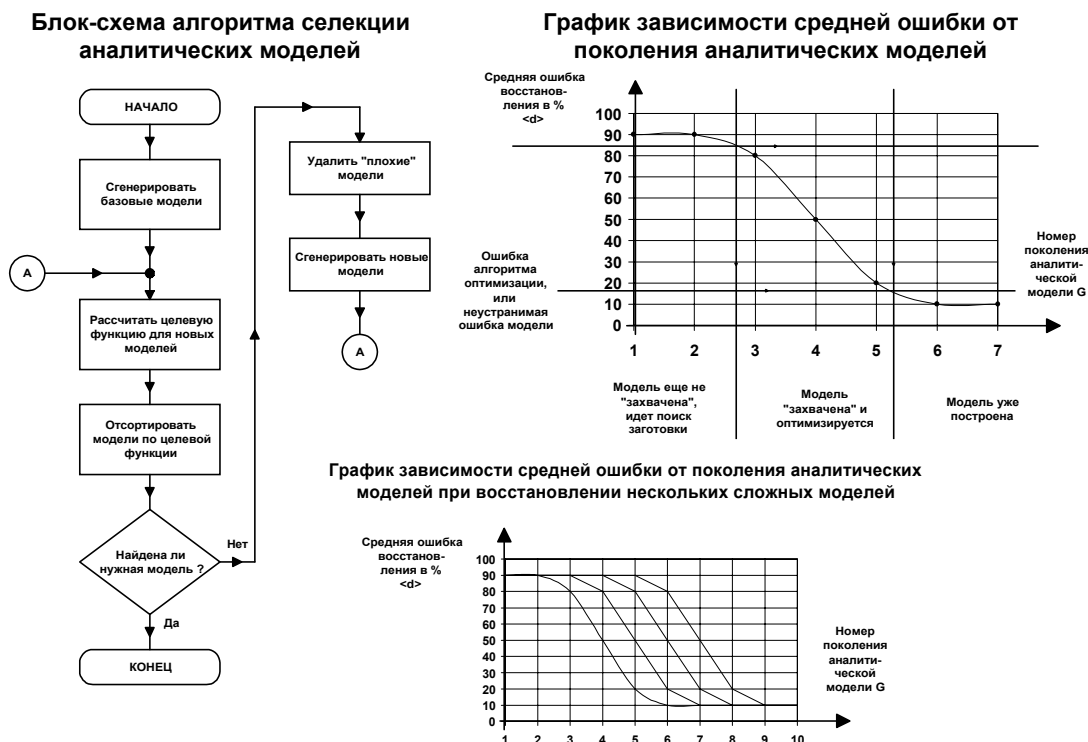


Рис. 3

ИСПОЛЬЗОВАНИЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА ДЛЯ НАХОЖДЕНИЯ ПАРАМЕТРОВ КОНСТРУИРУЕМЫХ МОДЕЛЕЙ

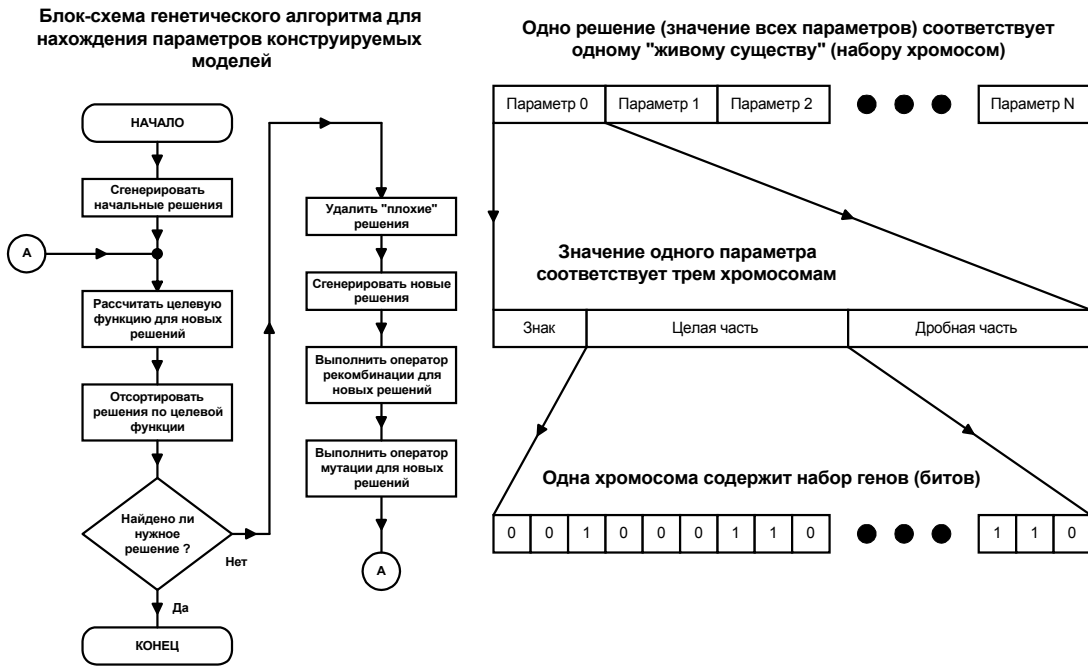


Рис. 4

ПРИМЕР ИСПОЛЬЗОВАНИЯ РАССМАТРИВАЕМОГО МЕТОДА ДЛЯ АППРОКСИМАЦИИ КРИВОЙ

Таблица с результатами конструирования функции одной переменной для аппроксимации кривой

Номер поколения аналитических моделей	Средняя ошибка аппроксимации	Аналитический вид модели лучшей для текущего поколения	Значения констант лучшей по поколению модели
0	848.856934	$(X)^*(C1)$	$C1 = -50.062500$
1	740.159973	$((X)-(C1))*((C2)-(X))$	$C1 = 6.519531, C2 = -144.281250$
2	366.103210	$(X)*((X)-(C1))*((X)*(C2))$	$C1 = -288.035400, C2 = 32.547363$
3	270.375610	$(X)*(((C1)-(X))*((X)-(C2)))-(X))$	$C1 = 9.799072, C2 = 16.315369$
4	166.799927	$(X)*((X)+(X)*(((C1)+(X))-((X)*(C2))))$	$C1 = 32.787582, C2 = 4.070465$
5	61.877922	$((X)+(X)*((X)*(((X)+(C1))-((X)*(C2)))))-(X))$	$C1 = 34.382141, C2 = 4.135788$

Графики кривых построенных функций и исходной функции

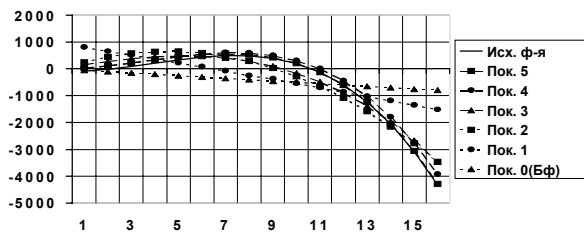


График зависимости средней ошибки моделирования от номера поколения

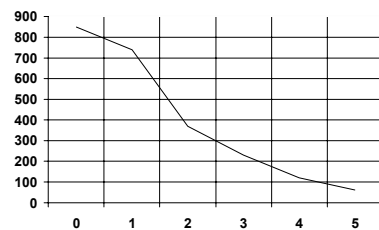


Рис. 5

Библиография

1. Ивахненко А.Г. Долгосрочное прогнозирование и управление сложными системами. - Киев: "Техника", 1975. - 312с.
2. Ивахненко А.Г. Индуктивный метод самоорганизации моделей сложных систем – Киев: "Наукова думка", 1982. – 296с.
3. Ивахненко А.Г. Моделирование сложных систем по экспериментальным данным. – М: "Радио и связь", 1987. – 119с.
4. Штейнберг Ш.Е. Идентификация в системах управления. – М: "Энергоатомиздат", 1978. – 288с.
5. Алгоритмы и программы восстановления зависимостей, Под ред. В.Н. Вапник. – М.:Наука, 1984 – 816с.
6. Курейчик В.М. Генетические алгоритмы. Монография. Таганрог.: Изд.-во ТРТУ, 1998, - 242с.
7. Кузнецов О.П., Адельсон-Вельский Г.М. Дискретная математика для инженера. – 2-е изд., перераб. И доп. – М.: "Энергоатомиздат", 1988. – 480с.



THE DEVELOPMENT OF ALGORITHMS AND PROGRAMS FOR THE CONSTRUCTION OF THE PROCESS ANALYTICAL DESCRIPTION BASED ON EVOLUTIONARY SEARCH

Shilov A.K., Shilov D.A., Selyankin V.V.

Taganrog State University of Radio Engineering
 347928, Russia, Taganrog, Nekrasovsky 44
 Tel: (86344) 310885,
 E-mail: t_sda@mail.ru

The report presents one method of self-organization of mathematical models. Such algorithms are widely spread in the systems of structure identification and complicated systems prognosis. The presented method can be applied for the curves and surfaces approximation. Which can be applied for the construction of the process analytical description and for the signal compression. Presented method of analytical model generation has four main features.

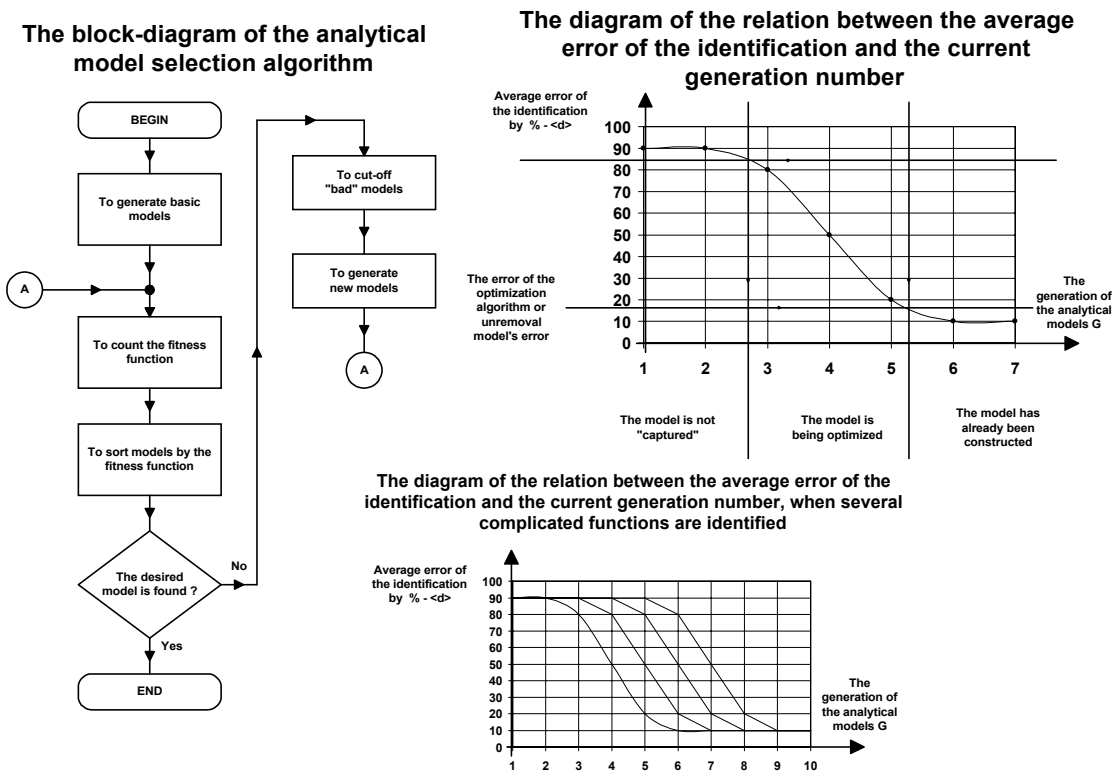


Fig. 1

Firstly, it is a complicated functions generation. It's based on two types of objects – basic functions and functional constructors.

Secondly, it is a decreasing of analytical model exhaust. For this purpose finite automata, formal grammars or tables with binary values can be used. In all cases the mathematical models generation process can be optimized. The other important method of decreasing exhaust of models uses the selection algorithms. The selection algorithm is one type of the wider class evolutionary algorithms. It simulates the process of artificial selection, which is used for breeding animals or plants with specific features. See fig. 1.

Thirdly, it is the searching for the functional parameters. To do this genetic algorithms can be used. They simulate natural biological evolution. In this case one chromosome contains all parameters of the model.

And the last feature is choosing some analytical models for the current part of source data. For this purpose the clustering algorithms can be applied.

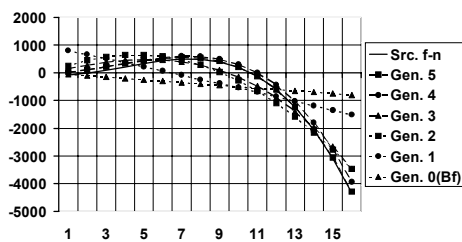
Now the algorithm has the precision of about ten percents in case when it is used for the function with five constant parameters and one variable parameter. See fig. 2. This algorithm is more flexible than its analogues, because it works with any shape of analytical description. And it can be used in the applications, where other algorithms have bad results.

THE EXAMPLE OF THE GIVEN METHOD USING FOR THE CURVE APPROXIMATION

The table with the results of one variable function construction which is used for the curve approximation

Analitical model's generation number	Middle error of approximation	Analitical shape of the generation's best model	Constant values of the generation's best model
0	848.856934	$((X)^*(C1))$	$C1 = -50.062500$
1	740.159973	$((X)-(C1))*((C2)-(X))$	$C1 = 6.519531, C2 = -144.281250$
2	366.103210	$((X)*((X)-(C1))-((X)*(C2))))$	$C1 = -288.035400, C2 = 32.547363$
3	270.375610	$((X)*(((C1)-(X))*(X)+(C2)))-(X))$	$C1 = 9.799072, C2 = 16.315369$
4	166.799927	$((X)*((X)+((X)*(((C1)+(X))-((X)*(C2))))))$	$C1 = 32.787582, C2 = 4.070465$
5	61.877922	$((X)+((X)*(((X)*((X)+(C1))-((X)*(C2)))-(X))))$	$C1 = 34.382141, C2 = 4.135788$

Curves diagrams of the constructed and source functions



The simulation error and generation number dependance diagram

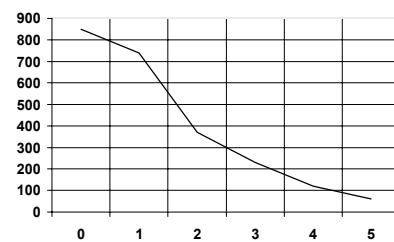


Fig. 2