

ОБНАРУЖЕНИЕ ЗАКОНОМЕРНОСТЕЙ ВЗАИМОДЕЙСТВИЯ ИОНОВ С ПОВЕРХНОСТЬЮ МАТЕРИАЛОВ ПО КОМБИНАТОРНОМУ АЛГОРИТМУ МГУА

Ивахненко А.Г.,¹ Савченко Е.А.,¹ Ивахненко Г.А.,² Надирадзе А.Б.,³ Рогов А.О.³

¹Международный центр информационных технологий и систем, Киев

²Национальный Институт Стратегических Исследований, Киев

³Московский авиационный институт, Москва

Введение

Вопросы взаимодействия ионов с поверхностью исследуются уже более 100 лет. За это время созданы десятки различных моделей, проведено огромное количество экспериментов. Тем не менее, эта проблема до сих пор остается актуальной и важной с практической точки зрения. Основная сложность ее решения состоит в том, что для адекватного описания процессов взаимодействия ионов с поверхностью необходимо учитывать большое количество факторов. Так, в настоящее время известно более 20 параметров, от которых зависит значение коэффициента распыления материалов [1], примерно те же параметры определяют значения коэффициентов аккомодации энергии и импульса [2]. Очевидно, что учесть все это многообразие в одной модели чрезвычайно сложно, особенно, если модель строится по принципу физической параметризации. Задачу построения модели можно трактовать, как задачу обнаружения закономерностей, содержащихся в выборке данных наблюдений сложного объекта, и рассматривать как один из примеров интерполяционных задач искусственного интеллекта.

Постановка задачи и первые алгоритмы ее решения были опубликованы еще в семидесятых годах прошлого столетия [3], но усовершенствование решений продолжается. Кроме того, продолжают развиваться алгоритмы вычисления дискриминантных функций распознающих систем, которые мало отличаются от алгоритмов обнаружения закономерностей [4]. Особенность данной работы состоит в том, что для обнаружения закономерностей применяются алгоритмы МГУА, основанные на самоорганизации моделей. Термин самоорганизация означает, что для определения моделей используются две выборки данных и, кроме того, они обрабатываются по очереди. Оценки коэффициентов полиномиальной модели получаются на одной выборке, а структура модели и число ее аргументов выбирается на другой выборке.

При таком способе обработки двух выборок, последовательно учитываются два критерия перебора множества моделей-кандидатов: внешний критерий ошибки модели $ER = RR_{A/A+B}$ и критерий ее смещения BS :

$$ER = RR_{A/A+B} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \rightarrow \min \quad (1)$$

$$BS = |RR_{A/A+B} - RR_{B/A+B}| \rightarrow \min \quad (2)$$

Сравнение выражений, используемых для расчета внешнего точностного критерия ER и критерия смещения модели BS, показывает, что при таком способе обработки двух выборок, последовательно учитываются два критерия перебора множества моделей-кандидатов: внешний критерий ошибки модели и критерий ее смещения. Смещение модели равно нулю, если модель оказывается на двух выборках одной и той же. Точностной критерий и критерий смещения не зависят друг от друга. Поэтому задача двухкритериального выбора модели, т.е. одновременного учета двух критериев, не возникает. В комбинаторных алгоритмах, разработанных в последнее время, сначала выбирается небольшое множество самых точных моделей, а затем применяется доопределение оптимальной модели по критерию смещения [5].

1. АЛГОРИТМ МГУА ДЛЯ ОБНАРУЖЕНИЯ ЗАКОНОМЕРНОСТЕЙ

Физические законы обладают свойством обобщения: они действуют в любое время и в любом месте пространства. Обнаружить эти законы можно по критерию смещения решений. По исходной выборке данных по комбинаторному алгоритму МГУА получается небольшое множество моделей с малым значением внешнего критерия ошибки ER. Из них нужно выбрать одну, наименее смещенную модель. Для этого выборка данных делится на две идентичные по размеру и дисперсии выборки А и В. Перебор постепенно усложняющихся структур моделей позволяет найти оптимальную, наименее смещенную модель.

Критерий смещения требует, чтобы модели равной сложности, полученные по методу наименьших квадратов (м.н.к.) на двух идентичных и независимых друг от друга выборках А и В на одном объекте, имели равные структуры и оценки коэффициентов. При этом будут равны и точностные характеристики моделей, полученных на А и В.

Поиск оптимальной модели состоит в расчете критерия смещения при постепенном изменении сложности структуры моделей. Расчет можно остановить, как только когда будет найден минимум критерия. Пример, приведенный ниже, показывает, как при помощи перебора вариантов можно выбрать полиномиальную модель с наименьшим значением критерия смещения.

1.1. Пример решения задачи обнаружения закономерности по критерию смещения

Пусть закономерность, подлежащая обнаружению по выборкам данных, представлена уравнением: $y = 20 + x_1 - 3x_5 + 5x_5$. Выборки А и В, представленные в таблице 1, отвечают этой зависимости.

Таблица 1. Исходные данные для обнаружения закономерности.

А											
N	x ₁	x ₂	x ₃	x ₄	x ₅	x ₆	x ₇	x ₈	x ₉	x ₁₀	y
1	4	2	7	10	1	5	8	9	6	3	8
2	2	10	3	4	7	8	5	1	9	6	48
3	7	4	9	3	5	1	6	10	2	8	25
4	3	1	5	8	2	10	9	6	7	4	18
5	10	8	4	1	3	6	2	8	5	7	33
6	8	3	2	6	10	4	7	5	1	9	72
7	9	5	8	9	6	2	3	7	4	1	35
8	5	7	6	2	4	9	1	3	8	10	27
9	6	9	1	7	8	3	4	2	10	5	63
10	1	6	10	5	9	7	10	4	3	2	36

В											
x ₁	x ₂	x ₃	x ₄	x ₅	x ₆	x ₇	x ₈	x ₉	x ₁₀	y	
3	2	1	10	9	8	7	6	5	4	65	
4	3	2	1	10	9	8	7	6	5	68	
5	4	3	2	1	10	9	8	7	6	21	
6	5	4	3	2	1	10	9	8	7	24	
7	6	5	4	3	2	1	10	9	8	27	
8	7	6	5	4	3	2	1	10	9	30	
9	8	7	6	5	4	3	2	1	10	33	
10	9	8	7	6	5	4	3	2	1	36	
1	10	9	8	7	6	5	4	3	2	29	
2	1	10	9	8	7	6	5	4	3	32	

Вычислительный эксперимент заключается в поиске этой закономерности, предполагая, что она неизвестна. Для решения этой задачи по данным выборок находим ряд постепенно усложняющихся моделей, и по м.н.к. находим оценки коэффициентов. Затем рассчитываем критерий смещения. Для сокращения объема, к перебору допускаются только аргументы, имеющие наибольшее значение модуля коэффициента корреляции с выходной величиной. В процессе постепенного усложнения моделей на седьмом ряду получено нулевое значение критерия смещения, указывающее, что эта модель заложена в выборках А и В. Полученная модель в точности совпадает с моделью, заложеной в выборки данных, а оценки ее коэффициентов пересчитываются по м.н.к. по всем имеющимся данным. На этом обнаружение закономерности заканчивается. На рис.1 показана зависимость точности и смещения от сложности модели.

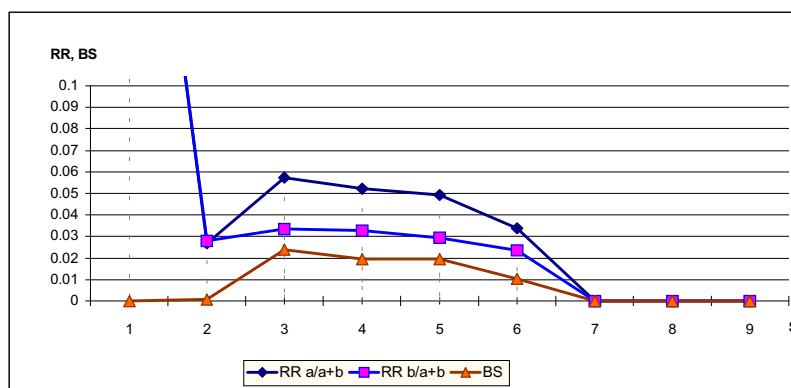


Рис.1. Зависимость точности и смещения от сложности модели.

2. ОСОБЕННОСТЬ АЛГОРИТМА ПЕРЕБОРА МОДЕЛЕЙ-КАНДИДАТОВ, ПОЗВОЛЯЮЩАЯ ОПРЕДЕЛИТЬ МНОЖЕСТВО НАИБОЛЕЕ ТОЧНЫХ МОДЕЛЕЙ

Как указывалось, внешний критерий ошибки модели и критерий смещения не связаны друг с другом. В зависимости оттого, что является целью выбора оптимальной модели, ее наименьшая ошибка или наименьшее смещение, может быть применен один из следующих вариантов дальнейшего развития алгоритма. Первый, – когда сначала по точности выбирается некоторое множество моделей, а затем по смещению из них выбирается лучшая. И второй, когда сначала по смещению выбирается некоторое множество наилучших моделей, а затем из них по точности выбирается одна, оптимальная.

При выборе множества наиболее точных моделей следует учитывать, что в области малой дисперсии помех, т.е. при точных данных, минимум критерия ошибки становится размытым и образуется интервал характеристики, на котором несколько моделей обладают почти одинаковой точностью. Именно эти модели и рекомендуются в качестве множества моделей, из которых выбирается одна наиболее несмещенная модель. Образование указанного интервала характеристики показано на рис.2., где сравниваются результаты дедуктивного расчета переборного процесса и экспериментальные данные.

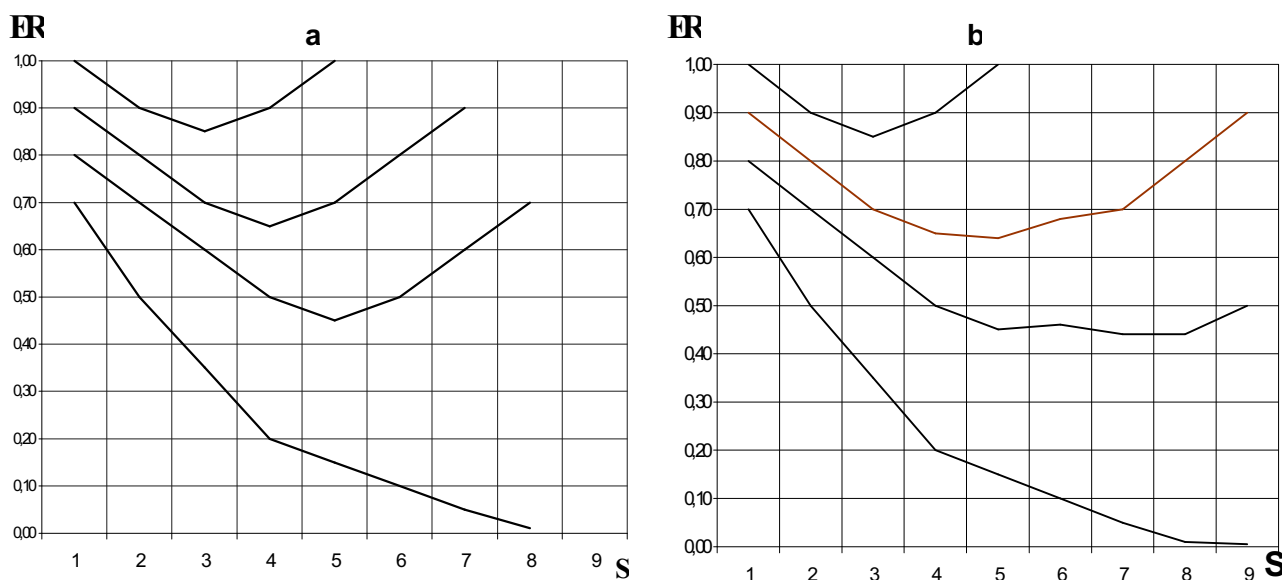


Рис.2. Сравнение результатов переборного процесса и экспериментальных данных.

2.2. Опыт исследования помехоустойчивости алгоритма обнаружения закономерностей

Вычислительный эксперимент состоял в том, что к выходной величине добавлялась знакопеременная аддитивная помеха. К выходной величине строк с четными номерами добавлялась определенная постоянная величина Δy , а из выходной величины данных с нечетными номерами строк, эта же величина вычиталась.

I. $\Delta y = +1, -1, +1, -1, \dots, +1, -1.$

II. $\Delta y = +2, -2, +2, -2, +2, -2, +2, -2, +2, -2.$

...

IV. $\Delta y = +5, -5, +5, -5, +5, -5, +5, -5, +5, -5.$

Результат эксперимента представлен на рис.3.

На рисунке показана зависимость ошибки оптимальной модели от ее сложности. Как видно из рисунка при постепенном увеличении помехи комбинаторный алгоритм изменяет только оценки коэффициентов, а структура моделей остается постоянной, содержащей восемь членов. Только при первом критическом значении помехи, равном $\Delta y = \pm 1,5$ происходит изменение структуры модели. Значение дисперсии помехи, при котором наблюдается первое изменение структуры модели, называется первым критическим значением дисперсии [6]. Это значение является границей между областью больших помех, где процесс перебора поддается не только экспериментальному исследованию, но и дедуктивному расчету, и областью малых помех, где результаты перебора можно получить только индуктивным методом.

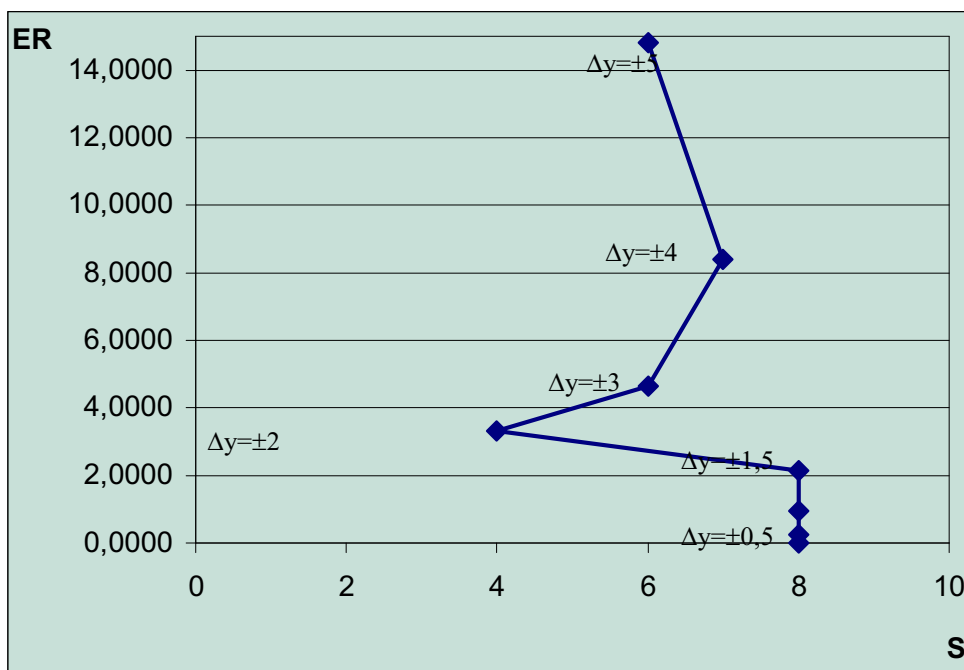


Рис.3. Результат экспериментального определения первого критического значения дисперсии перебора при аддитивной знакопеременной помехе $\pm\Delta y$.

При постепенном увеличении помехи достигается первое, критическое значение ее дисперсии, при котором оптимальной становится модель, имеющая только шесть аргументов. При дальнейшем повышении дисперсии помехи достигается второе критическое значение дисперсии, при котором амплитуда помехи равна $\pm\Delta y = 2$. В области малых помех наибольшее практическое значение имеет рассмотрение влияния помехи до первого критического значения дисперсии, где требуется применение добавочного определения оптимальной модели по критерию смещения.

2.3. Первое критическое значение дисперсии помех как граница области приложения дедуктивного подхода

Задача выбора оптимальной модели, соответствующей наименьшему значению внешнего критерия перебора множества моделей-кандидатов состоит в том, чтобы выбрать модель, дающую наименьшее значение критерия. Эта задача может быть решена как индуктивным, так и дедуктивным методом или подходом. При индуктивном подходе достаточно рассчитать значения критерия перебора для всех моделей и выбрать наименьшее значение. Для дедуктивного аналитического подхода нужно выбрать некоторую виртуальную схему перебора или другими словами принять некоторые ограничения. Например, в работе проф. Степашко В.С. [7] принимается, что помеха имеет нулевое среднее значение и действует аддитивно только на выходную переменную или, что одно и то же, сразу же на все аргументы. Объект хорошо описывается полиномиальным уравнением, причем минимум критерия единственный. При таких предположениях удастся рассчитать результат перебора. Проверка двух подходов на практических примерах показывает, что действительно дедуктивный подход дает те же результаты, что и индуктивный перебор всех моделей, но только в области значительных помех, т.е. при неточных данных. При дисперсии помех меньше первого критического

значения, зависимость критерия перебора от сложности модели образует небольшие колебания вокруг некоторого горизонтального участка (рис.2.).

Граница между областью, в которой применим дедуктивный аналитический подход и областью, где можно применить только индуктивный подход можно назвать границей между областью детерминированных и индетерминированных задач [8]. В области индетерминированных задач, называемой так же областью хаоса, алгоритмы МГУА успешно применяются только для интегральных характеристик, например для прогноза усредненных по интервалам значений переменных.

3. ПРИМЕР ПРИМЕНЕНИЯ АЛГОРИТМА МГУА ДЛЯ ОБНАРУЖЕНИЯ ЗАКОНОМЕРНОСТЕЙ, СОДЕРЖАЩИХСЯ В ВЫБОРКЕ ДАННЫХ НАБЛЮДЕНИЙ ВЗАИМОДЕЙСТВИЯ ИОНОВ С ПОВЕРХНОСТЬЮ МАТЕРИАЛОВ

Выборка исходных данных, нормированных по наибольшему значению для каждой переменной, представлена в таблице 2. В ней обозначены: x_1 – коэффициент распыления материала ионами ксенона при энергии 300 эВ и нормальном падении ионов [9], мг/К; x_2 – массовая плотность, г/см³; x_3 – молекулярный вес, а.е.м.; x_4 – температура сублимации, К; x_5 – теплоемкость Дж/моль/град; x_6 – энергия связи, эВ. Для краткости будем полагать, что выходная переменная искомым закономерности известна априори. Выходной будет служить переменная x_1 . Конечно, можно расширить поиск и считать выходными все переменные по очереди. Результатом поиска будут все закономерности с достаточно малым смещением.

Таблица 2. Исходные данные наблюдений взаимодействия ионов с поверхностью материалов.

	X1	x2	x3	X4	X5	x6
M	Y	Ro	Mw	Tsub	Cp	Uo
C	0,00681100	2,300	12,0000	4473	8,536	7,410
Be	0,02281800	1,848	9,0100	2744	16,440	3,480
Si	0,11722100	2,300	28,0855	3573	19,790	3,910
Ti	0,11756600	4,505	47,8800	3560	25,060	4,340
V	0,14095600	5,960	50,9400	3665	24,480	3,700
Cr	0,23101900	7,200	51,9600	2945	23,550	3,680
Nb	0,23225000	8,570	92,9000	5073	24,440	7,500
Al	0,24074400	2,689	26,9815	2793	24,350	3,260
Fe	0,31687600	7,870	55,8470	3145	24,980	4,150
Co	0,37866000	8,900	58,9332	3230	24,800	4,380
Mo	0,38376000	10,220	95,9400	4700	23,930	6,900
Ge	0,43554000	5,323	72,5900	3120	23,220	3,770
Cu	0,65659500	8,960	63,5460	2816	24,430	3,560
Ta	1,15879700	16,650	180,9400	5623	25,290	8,700
Pd	1,17062000	12,200	106,4200	3273	25,860	4,800
Mn	1,27194500	7,320	54,9380	2353	26,280	3,150
W	1,34088800	19,340	183,8500	5953	24,270	8,760
Pt	1,65818000	21,450	195,0800	4100	25,860	5,560
Ag	1,94148000	10,500	107,8600	2440	25,360	2,700
Au	2,36352000	19,320	196,9600	3150	25,400	3,920
Max	2,36352000	21,450	196,9600	5953	26,280	8,760

3.1. Первый вариант самоорганизации модели.

Вариант подробно описан в работе [6]. Учитывая свойства комбинаторного алгоритма МГУА для каждой модели должно быть отобрано не более 20 наиболее эффективных аргументов. В рассматриваемом примере их вышло ровно 20. Для этих аргументов-кандидатов, пользуясь комбинаторным алгоритмом МГУА, находим лучшие модели.

$$y=x1=f(x_2, x_3, x_4, x_5, x_6, x_2^2, x_2x_3, x_2x_4, x_2x_5, x_2x_6, x_3^2, x_3x_4, x_3x_5, x_3x_6, x_4^2, x_4x_5, x_4x_6, x_5^2, x_5x_6, x_6^2).$$

В примере таких моделей оказалось пять. Для них было рассчитано смещение по формуле (2).

На рис.4 и 5 приводятся результаты применения первого варианта выбора оптимальной модели.

Лучшей оказалась модель I:

$$y = 1,4343 - 5,856 x_6 + 0,4302 x_2 - 4,4347 x_2x_4 - 1,309 x_2x_5 + 7,891 x_2x_6 - 7,189 x_3x_4 + 4,781 x_3x_5 + 0,7226 x_4^2 - 0,58 x_4x_5 + 4,992 x_4x_6.$$

Показатели точности модели 1: ER = 0,0002728; BS = 3,55; MCC = 0,997.

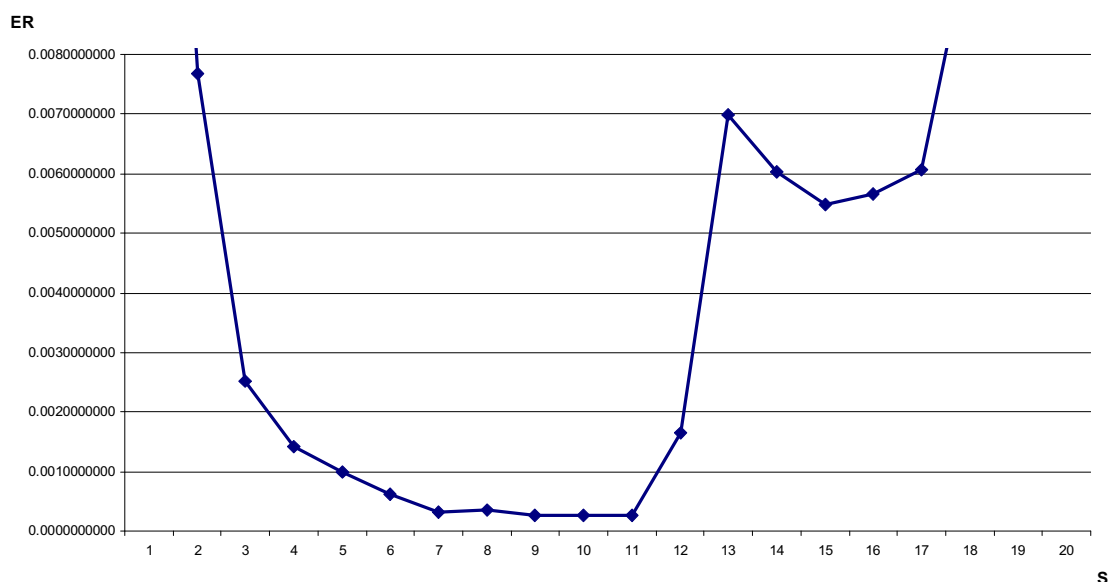


Рис.2. Зависимость внешнего критерия точности от сложности модели.

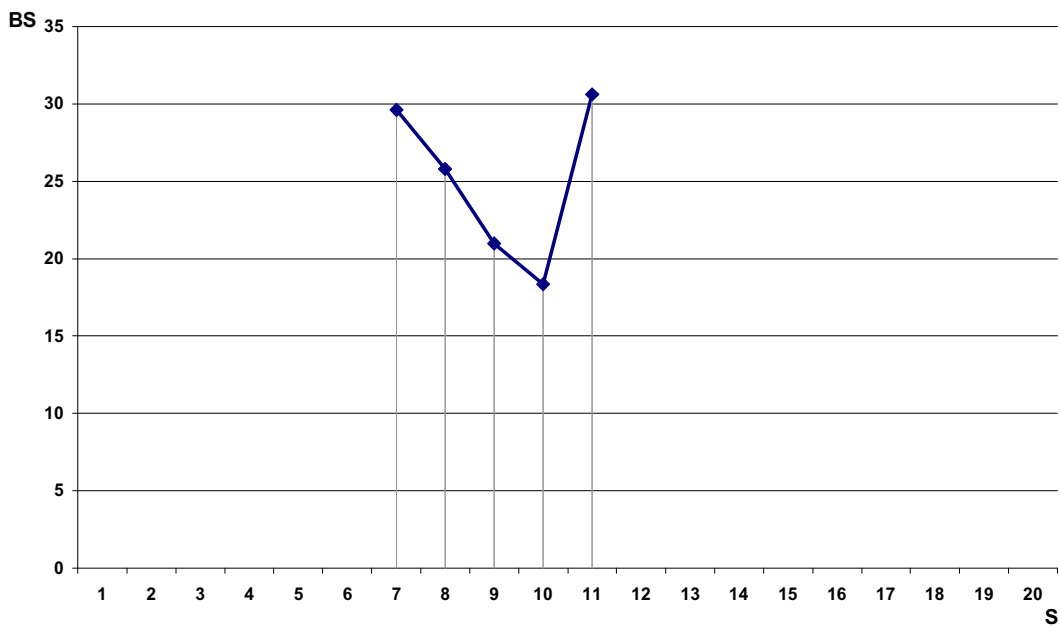


Рис.3. Доопределение модели по критерию смещения.

3.2. Второй вариант самоорганизации моделей.

Основное свойство физической закономерности состоит в том, что она обладает свойством обобщения на другие выборки данных. Это свойство можно оценить по величине критерия смещения BS. Смещение равно нулю, если зависимость на двух идентичных выборках А и В одна и та же. Поэтому сначала выбирается небольшое множество наименее смещенных моделей, а затем из них выбирается одна самая точная модель. На рис.6. показана зависимость смещения от сложности модели. В подмножество наиболее точных моделей выбираются те, для которых смещение оказалось наименьшим. Было выбрано семь моделей, лучших по смещению. Для них рассчитаны показатели точности.

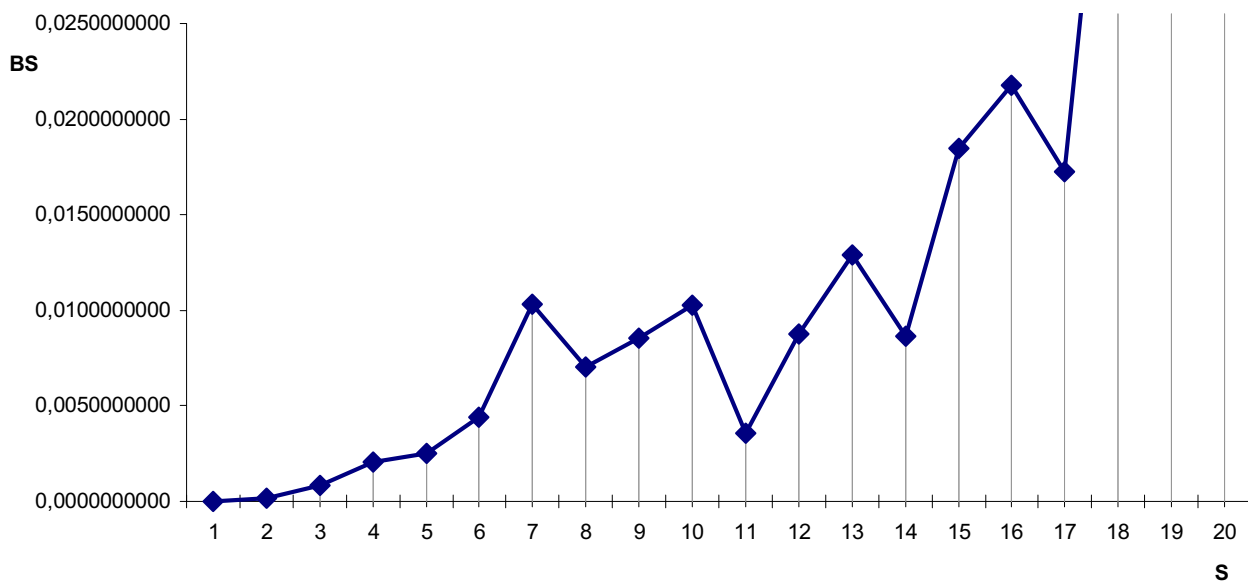


Рис.4. Зависимость смещения от сложности модели.

Рис.5. Зависимость точности от сложности моделей.

Лучшей оказалась модель II:

$$y = 0,3207 + 1,718 x_2 - 0,777 x_2^2 + 0,668 x_3x_4 - 1,388 x_3x_5.$$

Показатели точности модели 2: ER = 0,1705; BS = 0,002529; MCC = 0,911.

3.3. Выбор оптимальной модели по комбинированному критерию

Перебор моделей по критерию ошибки дал возможность выбора множества, состоящего из пяти наиболее точных моделей. Из этого множества моделей, по критерию смещения, выбрана модель I. Далее порядок перебора моделей был изменен. Сначала выбрано небольшое множество наиболее несмещенных моделей, из которых по критерию ошибки выбрана одна наиболее точная модель II. Из полученных таким образом моделей следует выбрать одну модель, оптимальную по двум критериям, т.е по величине комбинированному критерию, который вычисляется по формуле:

$$c = \sqrt{\lambda \cdot ER^2 + (1 - \lambda) \cdot BS^2} \rightarrow \min \quad (3)$$

Оптимальной назовем модель, в которой при вариации коэффициента λ в формуле (3) можно получить наименьшее значение c . При изменении λ от 0 до 1 для модели I получается меньшее значение комбинированного критерия. Следовательно, первая модель I является оптимальной, более точной, чем модель II. Алгоритм выбора оптимальной модели, в котором сначала находится множество достаточно точных моделей, называется алгоритмом с доопределением оптимальной по смещению [6].

4. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ВЫБОРКИ ДАННЫХ

4.1. Предварительная обработка данных при помощи корреляционного анализа информативной эффективности переменных

Модули коэффициента корреляции аргументов с выходной величиной показывают эффективность этих аргументов. К первичным шести аргументам были добавлены вторичные - парные ковариации первичных. Для них тоже были рассчитаны модули коэффициентов корреляции с выходной величиной. Для переменной x_1 наиболее влиятельным аргументом оказывается ковариация первичных аргументов x_2x_5 . Следующим по влиянию будет x_3x_5 и т.д. Корреляционная таблица также показывает пары переменных, между которыми слабая связь, например между x_1 и x_4 (таблица 3).

Таблица 3. Модули коэффициента корреляции аргументов с выходной величиной.

	x_2	x_3	x_4	x_5	x_6	x_2^2	x_2x_3	x_2x_4	x_2x_5	x_{2x}	x_3^2	x_3x_4	x_3x_5	x_3x_6	x_4^2	x_4x_5	x_4x_6	x_5^2	x_5x_6	x_6^2
$y=x_1$	0,810	0,813	0,005	0,464	0,021	0,781	0,789	0,594	0,882	0,543	0,788	0,591	0,825	0,539	0,042	0,243	0,054	0,512	0,237	0,054

4.2. Предварительная обработка выборки данных при помощи автоматической кластеризации

Строки выборки исходных данных можно рассматривать как множество координат точек в пространстве аргументов [18,19]. Автоматическая кластеризация позволяет разделить это множество на компактные группы точек, называемые кластерами. Каждый кластер объединяет в себе однородные реализации и потому может быть использован для выделения отдельной выборки. Для решения задач обнаружения закономерностей каждый кластер данных используется отдельно, что уменьшает ошибку и смещение моделей. Некоторые кластеры могут содержать слишком мало точек. В эти кластеры рекомендуется добавить координаты средних точек, равные среднеарифметическому координат всех точек каждого кластера. Для каждого кластера обнаруживается своя закономерность.

Заключение

С использованием алгоритма МГУА построены интерполяционные модели взаимодействия ионов с поверхностью, позволяющие прогнозировать значения коэффициентов распыления в зависимости от физических свойств облучаемых материалов. Установлено, что наиболее влиятельным аргументом моделей (среди рассмотренных) является ковариация первичных аргументов x_2x_5 и x_3x_5 или $\rho \cdot C_p$ и $M_w \cdot C_p$. Поскольку ранее влияние теплоемкости на коэффициенты распыления не рассматривалось, можно констатировать, что применение индуктивного перебора по внешним критериям позволило получить новое знание. Новое знание, полученное компьютером, связано с указанием переменной x_5 как величины, которая при достаточно большом x_2 влияет на коэффициент распыления.

Сравнение достигнутых значений критериев ошибки и смещения показывает, что более эффективным оказался вариант, при котором сначала выбирается некоторое множество наиболее точных моделей, а затем из них выбирается одна наименее смещенная модель.

Литература

1. Плешивцев Н.В., Бажин А.И., Физика воздействия ионных пучков на материалы, М.:Вузовская книга, 1998. – 392 с.
2. Баранцев Р.Г. Взаимодействие разреженных газов с обтекаемыми поверхностями. М.: Наука, 1975.
3. Загоруйко Н.Г. Прикладные методы анализа данных и знаний.- Новосибирск: Изд-во Ин-та математики, 1999.-270 с.
4. Круг Г.М., Круг О.Ю. Математический метод классификации древней керамики // Труды института археологии АН СССР. – Москва: Наука, 1965.-с.317 –323.
5. Ивахненко А.Г., Савченко Е.А., Ивахненко Г.А. Алгоритм МГУА для выбора оптимальной модели по внешнему критерию ошибки с доопределением по смещению модели и его применение в комитетах и нейросетях // Доклад на конференции МКИМ, Львов 2002.
6. Ивахненко А.Г., Степашко В.С.Помехоустойчивость моделирования. Киев: Наука, 1985 – 215 с.
7. Ивахненко А.Г., Зайченко Ю.П., Димитров В.Д. Принятие решений на основе самоорганизации. М: Сов. Радио, 1976. 280 стр.
8. Проблемы прикладной физики. Распыление твердых тел ионной бомбардировкой. Физическое распыление одноэлементных твердых тел/ Под ред. Р. Бериша: Пер. с англ./ Под ред. В.А. Молчанова. М.: 1984. - 336 с.
9. Ивахненко А.Г., Ивахненко Г.А. Проблемы дальнейшей разработки алгоритмов МГУА. Часть 1. // Проблемы управления и информатики, №6, 1999, с. 24-29.
10. Павловський М.А. Шлях України. К.: Техніка, 1996 –152 с.
11. Васильев В.И. Справочник по распознающим системам, 1975 – 258 с.
12. А.Г. Ивахненко. Долгосрочное прогнозирование и управление сложными системами. Киев.: Техника, 1975, 311 с.
13. Ивахненко А.Г., Ковалишин В.В., Тетко И.В., Луйк А.И., Ивахненко Г.А. Самоорганизация нейросетей с активными нейронами для предсказания активности химических соединений на основе алгоритма поиска аналогов, Проблемы управления и информатики, №.1, 1999, с.69-77.
14. Ivakhnenko A.G., Ivakhnenko G.A., Savchenko E.A., Gergely T. Twice multi-layered NN Self-Organization for Interpolation Problems of Artificial Intelligence Solution in the Case of Complete Absence of information about Small Part of Input Variables // Труды 15 Международной конференции по моделированию, Прага, июнь 2001 г.
15. Ивахненко А.Г., Ивахненко Г.А., Савченко Е.А., Вунш Д. Концепция последовательных алгоритмических приближений (спусков) к точному решению интерполяционных задач искусственного интеллекта.// Кибернетика и вычислительная техника, №6, 2000.
16. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. – М.: Статистика, 1973. – 392 с.
17. Мазуров В.Д., Метод комитетов в задачах оптимизации и классификации. М.: Наука, 1990, 250 с.
18. Ivakhnenko A.G., Ivakhnenko G.A. and Mueller J.-A. Self-Organization of Optimum Physical Clustering of Data Sample for a Weakened Description and Forecasting of Fuzzy Objects. \ Pattern Recognition and Image Analysis, v.3, N4, 1993, pp. 415-422.
19. Ивахненко А.Г., Аксенова Т.И. и др. Определение кластеров активности на поверхности молекул в области заданного химического действия.// Кибернетика и вычислительная техника, вып.118, 1998, стр.14-21.