

## ОСНОВНЫЕ ПРИНЦИПЫ МЕТОДА ГРУППОВОГО УЧЕТА АРГУМЕНТОВ И ЕГО ПЕРСПЕКТИВЫ

**Пидошва Е.А., Иващенко А.Б.**

Донецкий национальный технический университет

кафедра компьютерных систем мониторинга

E-mail: [katerinapidoshva@mail.ru](mailto:katerinapidoshva@mail.ru)

### *Аннотация*

*Пидошва Е.А., Иващенко А.Б. Основные принципы метода группового учета аргументов и его перспективы. Рассмотрены основные принципы метода группового учета аргументов. Показана полезность метода МГУА для различных областей науки. Установлено, что метод основан на индуктивных принципах - нахождение лучшего решения основано на переборе всевозможных вариантов. Определена принципиальная схема МГУА.*

### **Общая постановка проблемы**

Каждый человек в своей жизни в той или иной степени соприкасается с эмпирическими (описательными) социологическими исследованиями в качестве слушателя радио, читателя газет, журналов, научной литературы и других видов литературы. Возможно, и сам бывает вовлечён в такие исследования в качестве респондента, то есть источника исходной информации об изучаемых процессах и явлениях.

С помощью компьютерных имитационных экспериментов можно быстрее и проще решить ряд теоретических задач, по сравнению с использованием других математических методов. С использованием компьютерного имитационного моделирования осуществлено подтверждение некоторых положений классических и современных социологических теорий. Компьютерное имитационное моделирование позволило проверить некоторые положения социологических теорий (Г. Зиммель, Р. Мертон, Р. Уильямс) о динамике норм в индустриальных обществах.

### **Исследования**

Метод группового учета аргументов (МГУА) применяется в самых различных областях для анализа данных и отыскания знаний, прогнозирования и моделирования систем, оптимизации и распознавания образов. Индуктивные алгоритмы МГУА дают уникальную возможность автоматически находить взаимозависимости в данных, выбрать оптимальную структуру модели или сети, и увеличить точность существующих алгоритмов.

Этот подход самоорганизации моделей принципиально отличается от обычно используемых дедуктивных методов. Он основан на индуктивных принципах - нахождение лучшего решения основано на переборе всевозможных вариантов.

При помощи перебора различных решений подход индуктивного моделирования пытается минимизировать роль предпосылок автора о результатах моделирования. Компьютер сам находит структуру модели и законы, действующие в объекте. Он может быть использован при создании искусственного интеллекта как советчик для разрешения споров и при принятии решений [1].

Метод группового учета аргументов состоит из нескольких алгоритмов для решения разных задач. В него входят как параметрические, так и алгоритмы кластеризации, комплексирования аналогов, ребинаризации и вероятностные алгоритмы. Этот подход самоорганизации основан на переборе постепенно усложняющихся моделей и выборе наилучшего решения согласно минимуму внешнего критерия. В качестве базисных моделей

используются не только полиномы, но и также нелинейные, вероятностные функции или кластеризации.

Основной результат теории МГУА состоит в том, что при неточных зашумленных данных и коротких выборках, минимум критерия указывает нефизическую модель (решающее правило), точность которой выше и структура которой проще структуры полной физической модели. Перебор множества моделей-кандидатов по внешним критериям необходим только для нефизических моделей. При малых дисперсиях помех, рекомендуется дедуктивные алгоритмы, использующие обычные внутренние критерии перебора. При увеличении дисперсии помех приходится переходить к непараметрическим алгоритмам, использующим кластеризацию и поиск аналогов в предыстории, а для прогноза процессов применять эволюционное моделирование.

Заимствование алгоритмов переработки информации у природы является одной из основных идей кибернетики. «Гипотеза селекции» утверждает, что алгоритм массовой селекции растений или животных является оптимальным алгоритмом переработки информации в сложных задачах. При массовой селекции высевается некоторое количество семян. В результате опыления образуются сложные наследственные комбинации. Селекционеры выбирают некоторую часть растений, у которых интересующее их свойство выражено больше всего (эвристический критерий). Семена этих растений собирают и снова высевают для образования новых, еще более сложных комбинаций. Через несколько поколений селекция останавливается и ее результат является оптимальным. Если чрезмерно продолжать селекцию, то наступит «инцухт» — вырождение растений. Существует оптимальное число поколений и оптимальное количество семян, отбираемых в каждом из них.

Алгоритмы МГУА воспроизводят схему массовой селекции. В них есть генераторы усложняющихся из ряда в ряд комбинаций и пороговые самоотборы лучших из них. Так называемое «полное» описание объекта

$$\varphi = f(x_1, x_2, x_3, \dots, x_m),$$

где  $f$  — некоторая элементарная функция. Например, степенной полином заменяется несколькими рядами «частных» описаний:

$$1\text{-ряд селекции: } y_1 = f(x_1x_2), y_2 = f(x_1x_3), \dots, y_s = f(x_{m-1}x_m),$$

$$2\text{-й ряд селекции: } z_1 = f(y_1y_2), z_2 = f_2(y_1y_2), \dots, z_p = f(y_s - 1y_s), \text{ где } s = c^2, p = c^2_s \text{ и т. д.}$$

Входные аргументы и промежуточные переменные сопрягаются попарно, и сложность комбинаций на каждом ряду обработки информации возрастает (как при массовой селекции), пока не будет получена единственная модель оптимальной сложности.

Каждое частное описание является функцией только двух аргументов. Поэтому его коэффициенты легко определить по данным обучающей последовательности при малом числе узлов интерполяции. Исключая промежуточные переменные (если это удастся), можно получить «аналог» полного описания. Математика не запрещает обе эти операции. Например, по десяти узлам интерполяции можно получить в результате оценки коэффициентов полинома сотой степени и т. д.

Из ряда в ряд селекции пропускается только некоторое количество самых регулярных переменных. Степень регулярности оценивается по величине среднеквадратичной ошибки (средней для всех выбираемых в каждом поколении переменных или для одной самой точной переменной) на отдельной проверочной последовательности данных. Иногда в качестве показателя регулярности используется коэффициент корреляции.

Ряды селекции наращиваются до тех пор, пока регулярность повышается. Как только достигнут минимум ошибки, селекцию, во избежание «инцухта», следует остановить. Практически рекомендуется остановить селекцию даже несколько раньше достижения полного минимума, как только ошибка начинает падать слишком медленно. Это приводит к более простым и более достоверным уравнениям [2].

Решение практических задач и теории дизайна МГУА привели к развитию широкого спектра программных алгоритмов. Каждый из них соответствует определенным условиям его применения. Алгоритмы главным образом отличаются друг от друга расположением моделей кандидатов генераторных установок для данной основной функции, структуры модели комплексообразования и принимаются от внешних критериев. Выбор алгоритма зависит от специфики проблемы, уровня дисперсии шума, достаточности выборки данных, если выборка данных содержит только непрерывных данных.

Чаще всего используются [критерии](#) точности, дифференциальных или информативных типов. Работа алгоритмов МГУА имеет прямую аналогию с работой садовника при выборе новых гибридных растений.

Основные параметрические алгоритмы МГУА были разработаны для непрерывных переменных. Среди параметрических алгоритмов наиболее известными являются:

- основные комбинаторные ([COMBI](#)) алгоритмы (основаны на полном или сокращенном переборе постепенно сложных моделей и оценки их по внешнему критерию на отдельные части выборки данных);

- многослойный итерационный (МИА) алгоритм использует на каждом уровне процедуру сортировки по частичному описанию (его следует использовать, когда это необходимо для обработки большого числа переменных);

- целевой системный анализ, [OSA](#)-алгоритм - ключевой особенностью его является то, что он рассматривает не одно уравнение, а системы алгебраических или разностных уравнений (преимущество - что число регрессоров увеличивается, следовательно, к информации, содержащейся в данных образца используется лучше).

Непараметрические алгоритмы на примере:

- целевая компьютерная кластеризация (ОСС), которая работает с парами близко расположенных точек отбора проб;

- комплексирование аналогов ([АС](#)), которое использует множество аналогов, а не модели и кластеризации (рекомендуется для наиболее нечетких объектов);

- [вероятностный алгоритм](#), основанный на Многослойной Статистической Теории Решений. Рекомендуется для распознавания и прогнозирования бинарных объектов и для исходных данных управления, чтобы избежать возможных ошибок экспертов в нем.

Алгоритмы МГУА недавно применяются в оптимизации для решения проблемы нормативного прогнозирования и оптимального управления многомерными плохо определенными объектами. Алгоритмы МГУА применяются для оценки отклонения выходных переменных от их оптимальных значений.

Как уже отмечалось, алгоритмы МГУА были разработаны для непрерывных переменных. На практике, однако, образцы будут часто включать в себя дискретизованные переменные в небольшое число уровней или даже двоичные значения. Для расширения дискретизованных или бинарных переменных был разработан алгоритм МГУА гармонической дискретизации.

Наличие широкой гаммы алгоритмов МГУА прослеживается на то, что невозможно определить характеристики отдыха или контролируемых объектов точно заранее. Все вопросы, которые возникают во время процесса моделирования, должны быть решены путем сравнения критерия значения: что приводит к более глубокому [минимуму](#) основных внешних критериев. Таким образом, тип алгоритма выбирается объективно, по значению критерия дискриминации [3].

Принципиальная схема метода группового учета аргументов представлена далее на рисунке 1.

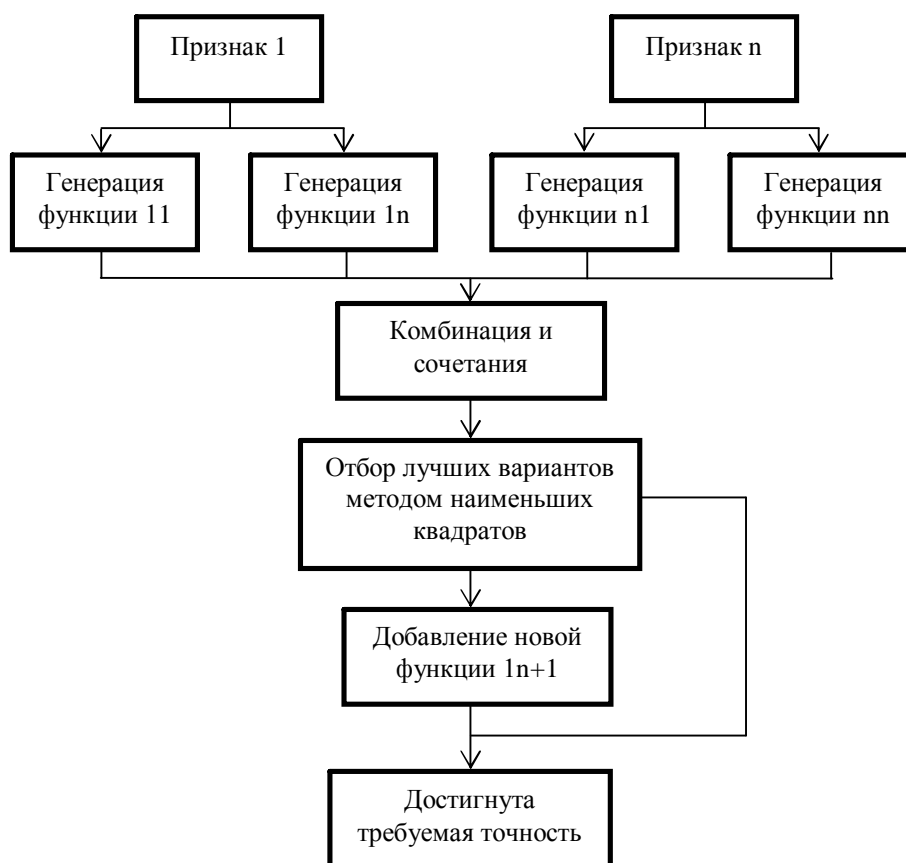


Рисунок 1 – Принципиальная схема МГУА

### Выводы

Направление МГУА может быть полезным потому что:

- находится оптимальная сложность структуры модели, адекватная уровню помех в выборке данных;
- количество слоев и нейронов в скрытых слоях, структура модели и другие оптимальные параметры нейросетей находятся автоматически;
- гарантируется нахождение наиболее точной или несмещенной модели - метод не пропускает наилучшего решения во время перебора всех вариантов;
- любые нелинейные функции или признаки, которые могут иметь влияние на выходную переменную используются как входные аргументы;
- переборные алгоритмы МГУА довольно просто запрограммировать;
- метод использует информацию непосредственно из выборки данных и минимизирует влияние априорных предположений автора о результатах моделирования;
- подход МГУА используется для повышения точности других алгоритмов моделирования.

### Література

1. Ивахненко А.Г. Долгосрочное прогнозирование и управление сложными системами. - К.: Техника, - 1975.
2. Ивахненко А.Г. Юрачковский Ю.П. Моделирование сложных систем по экспериментальным данным. – М.: Радио и связь, - 1987.
3. Ивахненко А.Г. Мюллер Й.А. Самоорганизация прогнозирующих моделей. К.: Техника, - 1985.