

**ПАРАЛЛЕЛЬНАЯ РЕАЛИЗАЦИЯ ГЕНЕТИЧЕСКОГО АЛГОРИТМА
ОПТИМИЗАЦИИ В ГИДРОЛОГИЧЕСКИХ МОДЕЛЯХ ТРАНСФОРМАЦИИ
СТОКА**

Сухотин М. А.

**научный руководитель канд. физ.-мат. наук, доц. Карпова Е.Д.
Сибирский федеральный университет**

В работе для восстановления параметров классической модели трансформации речного стока по данным наблюдений применен генетический алгоритм оптимизации и его параллельная реализация для высокопроизводительных вычислительных систем.

В рассмотренной модели производится вычисление расхода потока несжимаемой жидкости в выходном створе открытого русла при известном расходе потока во входном створе и известной передаточной функции. Рассматриваемая модель может быть записана в следующей форме:

$$Q(t) = \int_0^t q(\tau) f(t-\tau) d\tau, \quad (1)$$

где $q(\tau)$ – расход потока во входном створе в момент времени τ , $Q(t)$ – расход потока в выходном створе в момент времени t ($\text{м}^3/\text{с}$), $f(t)$ – передаточная функция. Выражение (1) представляет собой одну из форм интеграла Дюамеля.

Передаточная функция $f(t)$ в нашей модели представляет собой функцию вероятностного гамма-распределения, представленного в следующем виде:

$$f(t) = \frac{1}{\beta^\alpha \Gamma(\alpha)} t^{\alpha-1} e^{-\frac{t}{\beta}}. \quad (2)$$

Здесь α и β – параметры гамма-распределения. При этом $\alpha\beta$ – математическое ожидание распределения, а $\alpha\beta^2$ – его дисперсия.

Пусть в некоторые моменты времени $\{t_k\}_{k=1}^N$ известен наблюдаемый расход воды во входном створе русла $q_i = q(t_i)$. Соответственно, выражение (1) после дискретизации можно записать в следующем виде:

$$Q_j = \sum_{i=1}^j q_i \int_{t_j-t_i-\Delta t}^{t_j-t_i+\Delta t} f(\tau) d\tau. \quad (3)$$

Здесь Δt – половина среднего шага по времени. Для расчета по модели (2) – (3) необходимо знать параметры гамма-распределения α и β , которые в первом приближении можно подобрать исходя из характеристик потока.

Основными характеристиками потока, влияющими на решение задачи, являются средняя скорость потока v и параметр продольного рассеивания a . Параметры гамма-распределения являются функциями от v , a и значения длины исследуемого участка русла L , в случае (2) грубо их можно рассчитать по следующим формулам:

$$\alpha = \frac{L}{va^2}, \quad (4)$$

$$\beta = a^2. \quad (5)$$

Задача (2) – (5) является прямой задачей определения расхода по интегралу Дюамеля.

Для восстановления значений гамма-распределения α и β по данным наблюдений решалась следующая задача на оптимизацию в смысле наименьших квадратов.

Пусть имеется набор наблюдений за расходом потока в некоторые моменты времени $\{t_k\}_{k=1}^N$ во входном $\{q(t_k)\}_{k=1}^N$ и выходном $\{Q(t_k)\}_{k=1}^N$ створах. Определим параметры потока v и a , а, следовательно, и параметры гамма-распределения таким образом, чтобы доставлялся минимум функционалу:

$$\min_{v,a} \sum_{j=1}^N \left(\left(\sum_{i=1}^j q_i \int_{t_j-t_i-\Delta t}^{t_j-t_i+\Delta t} f(\tau) d\tau \right) - Q_j \right)^2. \quad (6)$$

Поскольку исследуемый бассейн в первые моменты времени еще не заполнен, прямой расчет по дискретному аналогу интеграла Дюамеля (2) в начале расчета всегда дает заниженный результат для первых нескольких моментов времени t_j . В результате первые несколько значений в функционале стоимости (6) будут давать большие ошибки и мешать оптимизации. В среднем можно ожидать удовлетворительных расчетов по формуле (6) только после $t_K \approx L/v$ шага по времени. Отбросим в (6) первые K слагаемых, оценив среднюю скорость потока по предоставленным натурным данным. В этом случае функционал стоимости (6) примет следующий вид:

$$\min_{v,a} \sum_{j=K}^N \left(\left(\sum_{i=1}^j q_i \int_{t_j-t_i-\Delta t}^{t_j-t_i+\Delta t} f(\tau) d\tau \right) - Q_j \right)^2.$$

Для решения поставленной оптимизационной задачи был выбран генетический алгоритм. Генетический алгоритм – это эвристический алгоритм поиска, используемый для решения задач оптимизации и моделирования путем случайного подбора, комбинирования и вариации искомых параметров с использованием механизмов, напоминающих биологическую эволюцию. Алгоритм является разновидностью эволюционных вычислений, с помощью которых решаются оптимизационные задачи с использованием методов естественной эволюции, таких как наследование, мутации, отбор и скрещивание.

Выбор метода решения обуславливается несколькими причинами. Во-первых, генетические алгоритмы обладают хорошим потенциальным параллелизмом. Во-вторых, поскольку генетические алгоритмы используют только значения самой функции, то не требуют никакой дополнительной информации (возможных оврагов, локальных минимумов и прочая), а также информации о гладкости оптимизируемой функции. В-третьих, применение генетических алгоритмов в гидрологических задачах пока имеет слабое распространение, тогда как они предоставляют широкие возможности, особенно в сочетании с вычислительными мощностями кластеров.

В общем виде ход алгоритма выглядит следующим образом:

1. Инициализация
2. Вычисление функции приспособленности
3. Цикл по количеству поколений:
 - 3.1 Генерация новой популяции
 - 3.2 Мутация
 - 3.3 Вычисление функции приспособленности

Генетические алгоритмы в ходе работы оперируют с хромосомами - бинарными строками, представляющими собой закодированные параметры оптимизации. Одна популяция состоит из большого числа хромосом, соответственно, необходимо в

каждом поколении много раз вычислять целевую функцию, а поскольку эти вычисления друг от друга не зависят, то именно этот этап алгоритма очень хорошо параллелизуется. К тому же, именно этап вычисления функции приспособленности является самым ресурсоемким в алгоритме, что делает его распараллеливание еще более необходимым. Функция приспособленности показывает, насколько хорошо данный индивид отвечает цели поставленной оптимизационной задачи.

На этапе генерации новой популяции с помощью оператора селекции выбираются два индивида, которые будут участвовать в скрещивании – обмене бинарных составляющих родительских хромосом. Выбор двух индивидов основан на значениях функции приспособленности для родительского поколения. Таким образом, индивиды с большим значением функции приспособленности имеют большую вероятность принять участие в формировании новой популяции.

На этапе мутации каждый ген (бинарная составляющая) каждого индивида с какой-то определенной алгоритмом вероятностью мутирует – меняет свое значение на противоположное. Мутация вводит в популяцию некоторое разнообразие и предупреждает потери, которые могли бы произойти вследствие исключения какого-нибудь значимого гена в результате скрещивания. К тому же, мутация обеспечивает поиск новых экстремумов, что повышает надежность нахождения глобального экстремума.

Для численного эксперимента были выбраны данные по расходу воды в створе Ивановской ГЭС (входной створ модели). В качестве замыкающего створа был выбран створ в 9,1 км ниже по течению. На рис. 1 отображен график расходов во входном створе специального пуска ГЭС, график наблюдаемых расходов в замыкающем створе и график расходов, рассчитанных по модели (3) после получения оптимизированных параметров гамма-распределения.

В расчетах $K=6$, из графика видно, что на первых шести шагах русло не заполнено и интеграл Дюамеля малоприспособлен для вычислений.

В работе были также оценены основные характеристики параллельной реализации генетического алгоритма оптимизации, который был реализован с помощью библиотеки MPI для языка Си. Расчеты проводились на кластере Института математики СФУ.

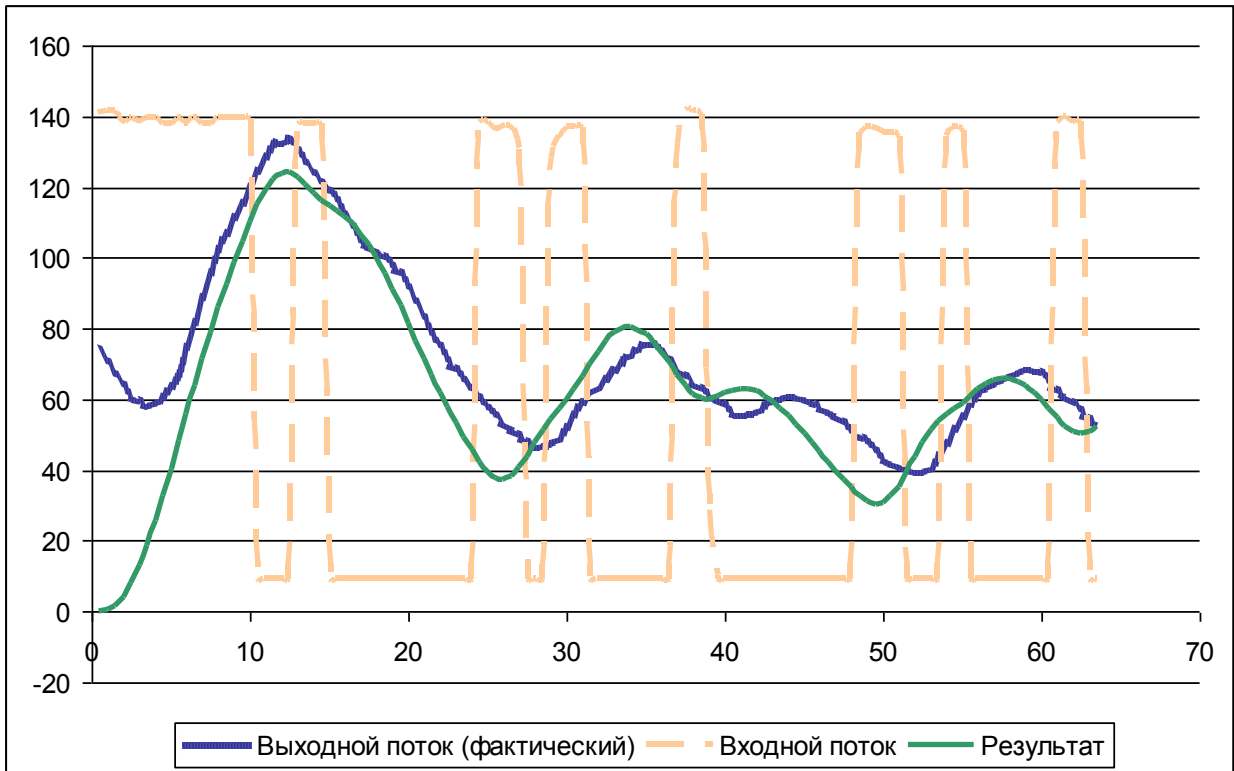


Рис. 1. Численный эксперимент по данным наблюдений на Ивановской ГЭС. По оси абсцисс отложено время в часах от начала специального попуска, по оси ординат – расход в m^3 / c

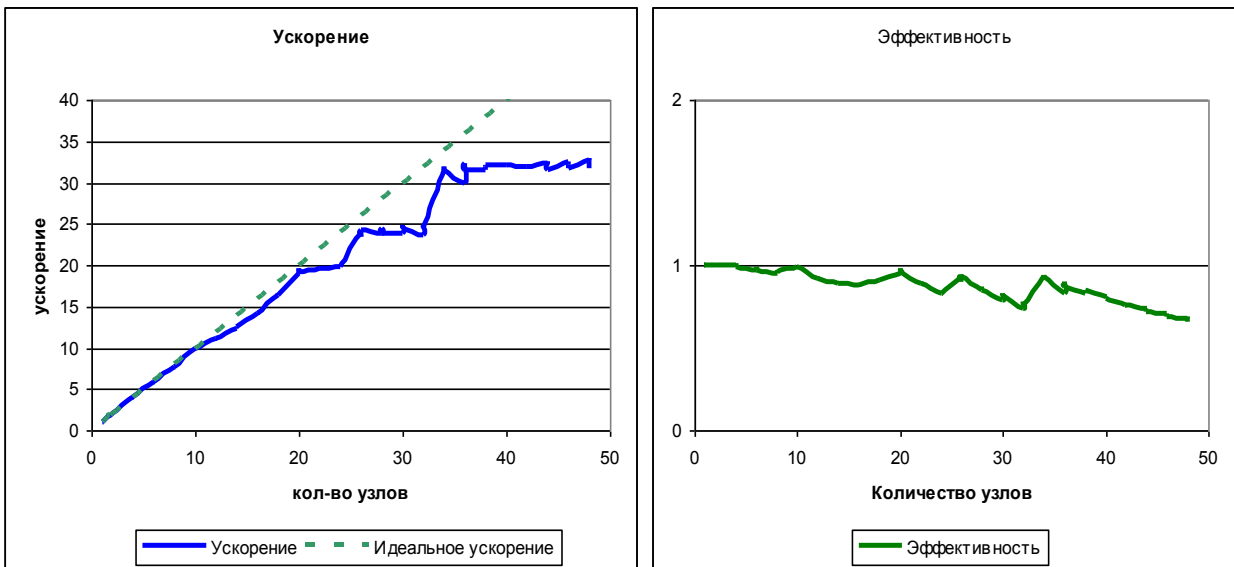


Рис. 2. Ускорение (слева) и эффективность (справа), полученные при расчетах на кластере ИМ СФУ

Ступенчатое поведение ускорения параллельного алгоритма объясняется его насыщаемостью. Этого можно избежать, перераспределив число поколений и количество особей в популяции. Влияние подобного перераспределения на конечный результат изучается в данный момент. Другой возможностью является совместное использование технологий распараллеливания задач OpenMP и MPI. Этот вариант также находится на стадии рассмотрения и тестирования.