

## ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

Tema:

# (Применение генетического алгоритма к задаче формирования инвестиционного портфеля»)

Выполнила:

студентка группы 381207 - 1

Пронина Елизавета Валерьевна

Научный руководитель:

кандидат технических наук, доцент

Неймарк Елена Александровна

# Введение: объект исследования, актуальность

- Классическая задача о ранце относится к числу широко известных задач дискретной оптимизации. Основные сферы применения находятся в областях планирования и управления экономическими, производственными и транспортными системами.
- Предлагаемый метод решения задачи о ранце генетический алгоритм. Данный алгоритм выбран в качестве объекта исследования.
- **Актуальность** исследования предопределена широкой распространенностью и важностью различных прикладных проблем, решаемым генетическим алгоритмом.

#### Цель работы

- ■Реализация генетического алгоритма для классической задачи о ранце.
- ■Подбор параметров ГА для наиболее эффективного поиска решения задачи формирования инвестиционного портфеля.

#### Практическое применение

В работе для исследования ГА рассмотрена экономическая задача о ранце - задача формирования инвестиционного портфеля, где в роли цены акции выступает вес, а выгодностью акции является стоимость в классической задачи о ранце.

Оптимизация инвестиционного портфеля с помощью генетических алгоритмов повышает качество инвестирования финансовых средств в виде надеждого сбережения капитала или получения максимального дохода при допустимом риске.



#### Постановка задачи о ранце

**Математическая:** Пусть дано п предметов. Для каждого і-го предмета задан вес  $w_i>0$  и ценность  $p_i>0$ , i=1,..., n. Задано ограничение на максимальный вес рюкзака -W. Каждый  $x_i$  может принимать только одно из двух значений:  $x_i=1$ , если i-й предмет попадает в рюкзак, или  $x_i=0$ , в противном случае. Требуется выбрать из заданного множества предметов набор с максимальной суммарной ценностью  $\sum_{i=1}^n p_i x_i$  при одновременном соблюдении ограничения на суммарный вес найденного набора  $\sum_{i=1}^n w_i x_i \le W$ .

$$\sum_{i=1}^{n} p_{i}x_{i} \to max$$

$$\sum_{i=1}^{n} w_{i}x_{i} \leq W$$

$$\sum_{i=1}^{n} w_{i} > W$$

$$x \in \{0,1\}, i = 1,2,...,n$$

$$p_{i} > 0, 0 < w_{i} \leq W, i = 1,2,...,n$$

#### Предлагаемый метод решения - ГА

▶ Генетические алгоритмы были предложены Джоном Генри Холландом в 1970 году и относятся к так называемым метаалгоритмам. Идея — составление алгоритмов поиска на основе биологической модели механизмов естественного отбора. Базовыми понятиями являются: популяция, отбор, мутация, скрещивание.

#### Плюсы ГА:

- 1/. Высокая скорость.
- 2. Независимость от класса исходной задачи.
- 3. Минусы:
- 4. Не гарантирует нахождение оптимального решения.



#### Программная реализация ГА

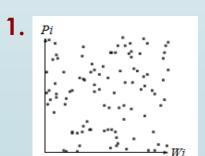
На языке С# разработана программа, решающая классическую задачу о ранце генетическим алгоритмом. В ней реализованы следующие операторы:

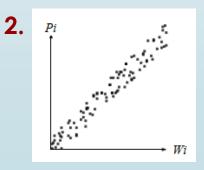
- 1. 2 оператора начальной популяции (алгоритм Данцига, случайный алгоритм)
- 2. 3 оператора кроссовера (одноточечный, двуточечный и однородный кроссоверы)
- 3. 4 оператора мутации (точечная мутация, сальтация, инверсия, транслокация)
- 4. 1 оператор обработки ограничений (метод штрафных функций + линейное динамическое масштабирование)
- 5. 2 оператора селекции (линейная ранговая селекция, бетта-турнир) + модификация поколения

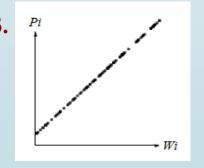
#### Классы тестовых задач

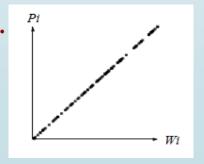
Экономическая 3ОР имеет особенности в виде корреляций между выгодностью и ценой акций, поэтому исследование проводится для 4 классов тестовых задач:

- 1. Задачи без корреляции (иллюстрируют примеры, когда вес и стоимость независимы).
- **2. Коррелированные задачи** (одни из наиболее реалистичных в управлении, возврат инвестиций, как правило, пропорционален вложенной сумме в некоторых небольших вориациях).
- 3. Сильнокоррелированные задачи (соответствуют реальной ситуации, где возвращается линейная функция от инвестиций).
- **4. Задачи с подсуммами** (отражают ситуацию, когда прибыль каждого элемента является линейной функцией веса).









 $P_{i}$  — цена i — того предмета,  $W_{i}$  — вес i — того предмета

#### Эксперимент: входные параметры

Мною было замоделировано поведение эволюционного генетического алгоритма при помощи программной реализации. Для подбора параметров ГА, влияющих на наиболее эффективный поиск оптимального решения был реализован алгоритм проведения «исследования ГА».

Для проведения исследования задаются входные параметры:

- Количество особей.
- Количество итераций (поколений).
- Так как перебираются все возможные комбинации операторов генетического алгоритма, учитывается задание значения β (для β – турнира, по умолчанию = 2).
- Количество экспериментов для конкретной задачи из класса тестовых задач. Один эксперимент - запуск программы с различными комбинациями операторов генетического алгоритма (всего 48 комбинаций = 48 запусков ГА для одного эксперимента).
- Количество запускаемых задач для каждого класса.

#### Эксперимент: количество запусков

#### Для одного запуска генетического алгоритма взяты следующие параметры:

- 30 особей
- 40 поколений
- β = 14 (для бетта турнирной селекции)

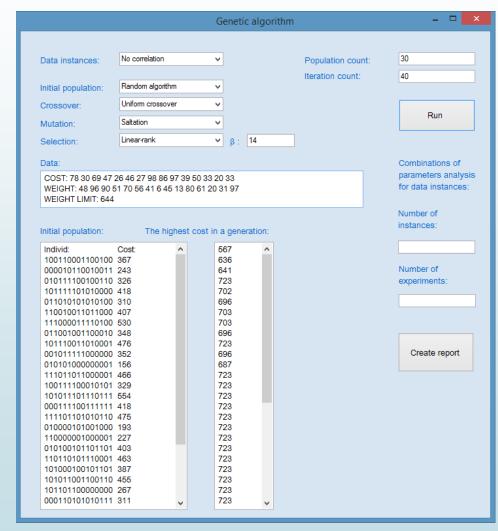
#### Всего проведено 600 экспериментов (28800 запусков ГА), из них:

- 30 экспериментов на 1 задачу
- По 5 задач на конкретный класс тестовых задач.
- Всего 4 класса тестовых задач

Для генерации задачи диапазон значений веса и цены - R = [0, ..., 100].

Каждая задача прорешена программно методом полного

перебора, что бы найти точное решение для исследования.



#### Эксперимент: вывод результатов

Лист №1 Эксперимент №1

Лист №2 Эксперимент №2

Книга №1 Задача №1

Лист №30 Эксперимент №30

Лист №1 Эксперимент №1

Лист №2 Эксперимент №2

Книга №5 Задача №5

Лист №30 Эксперимент №30

- Значения ф-й приспособленности для кажд. поколения (40 шт)
- Максимальное значение ф-и присп.
- Кол-во итераций до лучшего решения

Лист №31 Суммарный отчёт по всем экспериментам

Лист №31 Суммарный отчёт по всем экспериментам

- Макс. значение ф-и приспособленности
- Вероятность получения лучшей особи
- Среднее значение количества итераций до получения лучшей особи

Книга №6 Суммарный отчёт по всем задачам одного класса тестовых задач

- Мин. отклонение значения функции приспособленности от точного решения %
- Средняя вероятность получения лучшей особи %
- Среднее значение количества итераций % от наибольшего среднего количества итераций до получения лучшей особи (средняя скорость до наилучшего решения)

#### Анализ результатов

Из полученых результатов исследования для каждого класса тестовых задач выбрано по 5 эффективных комбинаций. Критериями отбора являлись:

- Вероятность нахождения глобального оптимума
- Скорость нахождения глобального оптимума

Для всех классов тестовых задач выбраны три эффективные комбинации. Критерий отбора - количество классов, содержащих комбинацию.

Данные результаты так же проанализированы и сделаны выводы, подтверждающие теоретический материал:

- **Относительно ГА:** Во всех оптимальных комбинациях в качестве оператора селекции выступает линейная ранговая схема. Это подтверждает теорию о том, что данная схема предотвращает преждевременную сходимость и приводит к наиболее лучшему решению.
- Относительно классов тестовых задач: Для задач с корреляцией и с сильной корреляцией вероятность нахождения глобального оптимума не превышает 22,67%, что подтверждает теорию о труднорешаемости таких задач, напротив, для задач без корряции и с подсуммами вероятность нахождения глобального оптимума достигает 72% и 84,67%, соответственно, что говорит о менее трудном нахождении их решения.

#### Задачи без корреляции

1	2
L	

	Алгоритм				
	Данцига	Случайный	Алгоритм	Случайный	Случайный
	Однородный	алгоритм	Данцига	алгоритм	алгоритм
	кроссовер	Однородный	Однородный	Однородный	Однородный
	Точечная	кроссовер	кроссовер	кроссовер	кроссовер
	мутация	Транслокация	Инверсия	Сальтация	Инверсия
	Линейная	Линейная	Линейная	Линейная	Линейная
	ранговая	ранговая	ранговая	ранговая	ранговая
	схема	схема	схема	схема	схема
Вероятность %	72%	69,33%	69,33%	69,33%	68,67%
Среднее кол-во итераций %	17,95%	16,93%	17,17%	17,4%	17,62%

#### Задачи с корреляцией

	Случайный				
	алгоритм	Алгоритм	Алгоритм	Алгоритм	Случайный
	Однородный	Данцига	Данцига	Данцига	алгоритм
	кроссовер	Однородный	Однородный	Однородный	Однородный
	Точечная	кроссовер	кроссовер	кроссовер	кроссовер
	мутация	Сальтация	Инверсия	Транслокация	Сальтация
	Линейная	Линейная	Линейная	Линейная	Линейная
	ранговая	ранговая	ранговая	ранговая	ранговая
	схема	схема	схема	схема	схема
Вероятность %	22,67%	20,67%	20,66%	18,67%	17,33%
Среднее кол-во итераций %	17,57%	16,6%	18,32%	17,32%	14,82%

1	4

				Случайный	
	Алгоритм	Алгоритм	Случайный	алгоритм	Случайный
	Данцига	Данцига	алгоритм	Двуточечный	алгоритм
	Двуточечный	Одноточечный	Однородный	кроссовер	Однородный
	кроссовер	кроссовер	кроссовер	Точечная	кроссовер
	Инверсия	Инверсия	Сальтация	мутация	Транслокация
	Линейная	Линейная	Линейная	Линейная	Линейная
	ранговая	ранговая	ранговая	ранговая	ранговая
	схема	схема	схема	схема	схема
Вероятность %	22,67%	17,33%	16%	16%	15,33%
Среднее кол-во итераций %	15,62%	15,72%	14,03%	17,13%	13,18%

#### Задачи с подсуммами

			Случайный		
	Алгоритм	Алгоритм	алгоритм	Случайный	Случайный
	Данцига	Данцига	Однородный	алгоритм	алгоритм
	Однородный	Однородный	кроссовер	Однородный	Однородный
	кроссовер	кроссовер	Точечная	кроссовер	кроссовер
	Сальтация	Инверсия	мутация	Сальтация	Транслокация
	Линейная	Линейная	Линейная	Линейная	Линейная
	ранговая	ранговая	ранговая	ранговая	ранговая
	схема	схема	схема	схема	схема
Вероятность %	84,67%	80,67%	78%	76,66%	74%
Среднее кол-во итераций %	13,1%	13,78%	14,62%	13,23%	13,01%

#### По всем классам задач

	Случайный алгоритм Однородный кроссовер Сальтация Линейная ранговая схема	Алгоритм Данцига Однородный кроссовер Инверсия Линейная ранговая схема	Случайный алгоритм Однородный кроссовер Сальтация Линейная ранговая схема
Колличество классов, содержащих комбинацию	4	3	3

#### Заключение

**Практическая значимость исследования** состоит в том, что подобраные эффективные наборы параметров генетического алгоритма для задачи формирования инвестиционного портфеля ускорят процесс и дадут наиболее точное решение для нахождения максимальной выгодности акций по ограниченной цене.

Так как проблемы, решаемые генетическим алгоритмом, являются актуальными и по сей день, в дальнейшем можно продолжить его исследование исходя из других свойств:

- Метод кодировки решений.
- Настройки параметров, такие как:
  - ▶ количество особей,
  - количество итераций.
- Частный критерий успеха.

### Спасибо за внимание!