

Применение методов Монте-Карло для решения многокритериальных задач

Выпускная квалификационная работа

Бакшинская Екатерина Олеговна, гр. 14.Б02-мм

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

Научный руководитель: к.ф.-м.н., доцент Шпилев П.В.

Рецензент: к.ф.-м.н., доцент Алексеева Н.П.



13 июня 2018 г.

Данная работа посвящена развитию подхода, основанного на применении *скрытых марковских моделей (СММ)* к решению многокритериальных задач теории игр.

Типы игр:

- кооперативные и некооперативные;
- с нулевой суммой и ненулевой суммой;
- конечные и бесконечные;
- с полной и неполной информацией.

Данная работа посвящена развитию подхода, основанного на применении *скрытых марковских моделей (СММ)* к решению многокритериальных задач теории игр.

Цель работы:

Разработка алгоритма для определения оптимальных стратегий в играх с неполной информацией.

Идея:

- Использование СММ для описания состояний игры в различные моменты времени.
- Представление поведения игроков как набор состояний.

Скрытая марковская модель с параметрами $\lambda = (A, B, \pi)$ описывается следующим образом:

- Конечное множество скрытых состояний $S = \{s_1, \dots, s_n\}$.
- Вероятности переходов для состояний, матрица A :

$$A = \{a_{ij}\}, \quad a_{ij} = P(S_t = j | S_{t-1} = i).$$

- Конечное множество наблюдений $O = \{o_1, \dots, o_k\}$.
- Вероятности переходов из состояний в наблюдения, матрица B :

$$B = \{b_{il}\}, \quad b_{il} = P(O_t = l | S_t = i).$$

- Вектор π задающий вероятность того или иного состояния в начальный момент времени:

$$\pi = \{\pi_i\}, \quad \pi_i = P(S_1 = i).$$

Постановка задачи

Задача: Для игрока i узнать вероятность, что $O_{t_{l+1}} = o_j$, при $1 \leq j \leq k$, учитывая последовательность наблюдений $\tilde{\mathbf{O}} = \{O_1, \dots, O_t\}$.

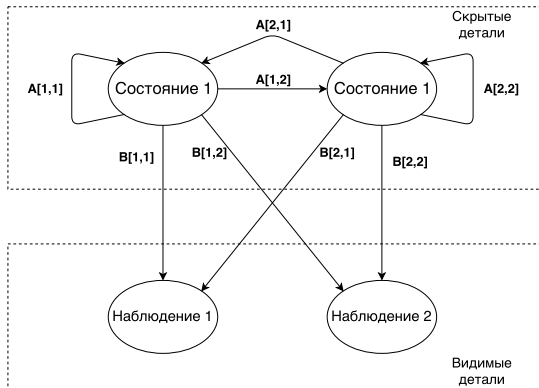


Рис. 1: Пример СММ с двумя состояниями и двумя наблюдениями

Входные данные:

- 2 игрока;
- множество доступных игрокам действий (наблюдений):
 $\mathbf{O} = \{o_1, \dots, o_k\}$;
- множество состояний 1-го игрока: $\mathbf{S} = \{s_1, \dots, s_n\}$;
- матрицы выплат для каждого из состояний $\{U_1, \dots, U_n\}$ размерами $k \times k$.

Замечание: поведение 1-го игрока описывается с помощью СММ с параметрами $\lambda = (\mathbf{A}, \mathbf{B}, \pi)$.

Результат: на каждом шаге игры алгоритм выдает наиболее вероятную последовательность скрытых состояний 1-го игрока со всеми параметрами его СММ.

Алгоритм Баума-Велша позволяет оценить неизвестные параметры СММ.

Входные данные:

- последовательность наблюдений $\tilde{\mathbf{O}}$;
- СММ с параметрами $\lambda = (\mathbf{A}^*, \mathbf{B}, \pi^*)$, где матрица \mathbf{A}^* и вектор π^* вероятностные произвольные.

Результат: СММ с параметрами $\lambda^* = \max_{\lambda} P(\tilde{\mathbf{O}}|\lambda)$.

Алгоритм Витерби позволяет сделать предположение о последовательности скрытых состояний СММ на основе последовательности наблюдений.

Входные данные:

- последовательность наблюдений $\tilde{\mathbf{O}}$;
- СММ с параметрами λ^* .

Результат: наиболее вероятная последовательность скрытых состояний противника $\tilde{\mathbf{S}}$.

Пошаговая схема алгоритма решения

- 1 вычисляем элементы матрицы \mathbf{B} , находя смешанные стратегии для первого игрока по матрицам выплат;
- 2 первые 3 хода игры второй игрок выбирает случайным образом;
- 3 формируется последовательность наблюдений (ходов первого игрока): $\tilde{\mathbf{O}} = \{\mathbf{O}_1, \mathbf{O}_2, \mathbf{O}_3\}$;
- 4 используем алгоритм Баума–Велша для вывода оценок параметров \mathbf{A} и π ;
- 5 находим последовательность скрытых состояний 1-го игрока $\tilde{\mathbf{S}}$, используя алгоритм Витерби.
- 6 шаги (4), (5) \Rightarrow имеем все сведения о 1-м игроке \Rightarrow 2-й игрок может выбрать ход максимизирующий его выигрыш;
- 7 добавляем результат последнего хода 1-го игрока к цепочке наблюдений $\tilde{\mathbf{O}}$;
- 8 повторяем шаги (4)–(7) после каждого хода до окончания игры.

Пример игры для тестирования

Игра в теннис, описанная в (“Using HMM in Strategic Games”, 2013):

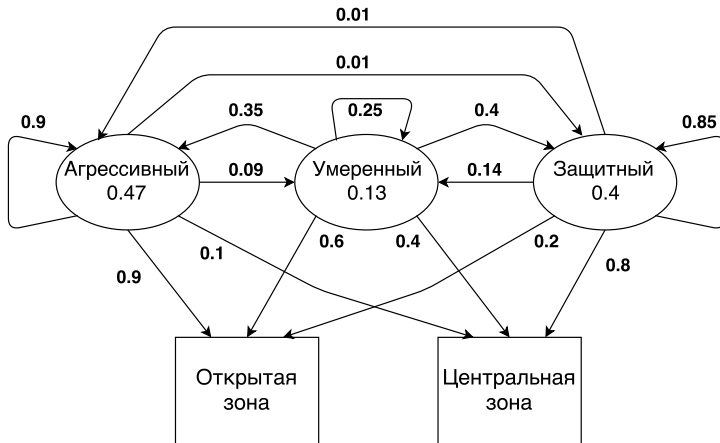


Рис. 2: СММ подающего игрока (сценарий “агрессивный игрок”)

$$\mathbf{B} = \{b_{il}\}, \quad b_{il} = P(\mathbf{O}_t = l | \mathbf{S}_t = i).$$

Таблица 1: Матрица \mathbf{B}_1

	Открытая зона	Центральная зона
Агрессивный игрок	0.61	0.39
Умеренный игрок	0.54	0.46
Защитный игрок	0.64	0.36

Таблица 2: Матрица \mathbf{B}_2 из статьи “Using HMM in Strategic Games”, 2013

	Открытая зона	Центральная зона
Агрессивный игрок	0.9	0.1
Умеренный игрок	0.6	0.4
Защитный игрок	0.2	0.8

Сравнение результатов по данным из статьи “Using HMM in Strategic Games”, 2013 (с матрицей B_2).

Таблица 3: Сравнение результатов (сценарий агрессивный игрок)

	Наш алгоритм	Алгоритм, предложенный в статье
Количество побед	94%	78%

Таблица 4: Сравнение результатов (сценарий защитный игрок)

	Наш алгоритм	Алгоритм, предложенный в статье
Количество побед	92%	71%

Данные из статьи “Using HMM in Strategic Games”, 2013.

Сценарий агрессивный игрок:

Количество отраженных ударов: 74,53%.

Доверительный интервал: (71,42%; 76,24%).

Сценарий защитный игрок:

Количество отраженных ударов: 73,67%.

Доверительный интервал: (71,18%; 76,64%).

Таблица 5: Результаты тестирования для B_1

	Результат, посчитанный по матрицам выплат	Количество отраженных ударов
Сценарий агрессивный игрок	62%	59.24%
Сценарий защитный игрок	62%	61.34%

Доверительные интервалы:

Для сценария агрессивного игрока: (64.5%; 75.42%).

Для сценария защитного игрока: (63.89%; 74.18%)

Случай 1: Неравномерные матрицы A и B .

Случай 2: Неравномерная матрица A и почти равномерная B .

Случай 3: Почти равномерная матрица A и неравномерная B .

Случай 4: Почти равномерные матрицы A и B .

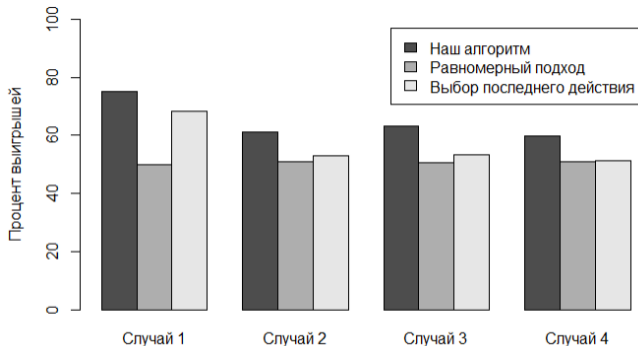


Рис. 3: Сравнение результатов нашего алгоритма с другими подходами к решению поставленной задачи

Модель игры “Камень, ножницы, бумага”:

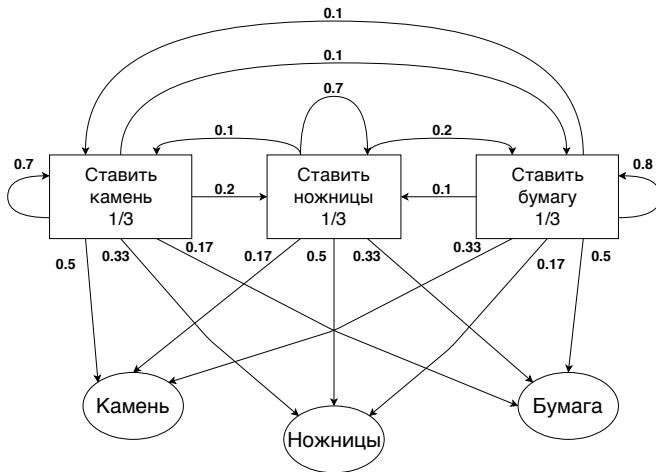


Рис. 4: Оригинальная СММ игры “Камень, ножницы, бумага”

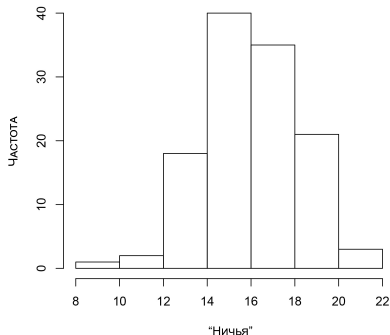
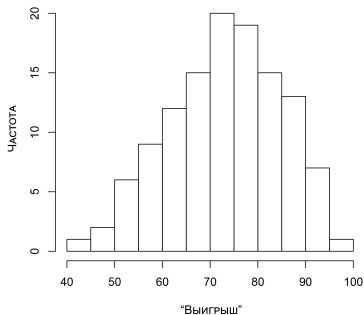
Результаты, полученные при тестировании предложенной модели решения на игре “Камень, ножницы, бумага”:

Таблица 6: Результаты в игре “Камень, ножницы, бумага”

	Количество “выигрышей”	Количество ситуаций “ничья”
Результат	74.3%	15.2%

Тестирование на произвольных данных игры “камень, ножницы, бумага”

Результаты 100 запусков алгоритма:



Среднее по количеству выигрышей: 72.4%.

Среднее для ситуации “ничья”: 15.7%.

Основные результаты:

- ❶ Использованы СММ (скрытые марковские модели) в качестве инструмента описания данного процесса (стохастической игры).
- ❷ Представлен алгоритм решения частных случаев стохастической игры с отсутствием полной информации.
- ❸ Использованы алгоритмы Баума-Велша и Витерби как инструмент решения задач, описанных СММ.
- ❹ В практической части:
 - Алгоритм был протестирован на гипотетическом примере игры в “теннис”, описанном в “Using HMM in Strategic Games”, 2013.
 - Было проведено сравнение результатов работы предложенного алгоритма с результатами применения других подходов к решению поставленной задачи.
 - Предложенный алгоритм был протестирован на произвольных данных другой игры со схожими характеристиками: “камень, ножницы, бумага”.