

Научная статья
УДК 794.24:004.42
doi: 10.34822/1999-7604-2022-3-14-24

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И АНАЛИЗ ДЛЯ ОПРЕДЕЛЕНИЯ СИЛЫ ИГРОКА СЁГИ

Ольга Павловна Бобровская^{1✉}, Светлана Александровна Лысенкова²

^{1,2} Сургутский государственный университет, Сургут, Россия

¹ o-bobrovskaya@mail.ru ✉, <https://orcid.org/0000-0001-7045-9085>

² lysenkova_sa@surgu.ru, <https://orcid.org/0000-0003-1007-7610>

Аннотация. Рассматриваются численные оценки позиций фигур на доске в партии сёги, получаемые от компьютерной программы. С использованием этих оценок предполагается рассчитать силу игрока, для чего требуется собрать данные, рассмотреть и повторить существующие подходы, предложить и реализовать свой подход, сравнить результаты. Проводится повторение подхода Феррейры, в котором на основании распределения выигрыша двух игроков рассчитывается разница их силы. При реализации подхода Ямаситы ищется зависимость между рейтингом Эло и средним ухудшением оценки позиции в результате плохого хода. Предлагается собственный подход к оценке силы игрока. Для этого рассматривается зависимость процента побед, хороших и плохих ходов, разницы оценки реальных и идеальных ходов и среднего улучшения позиции от рейтинга Эло. Предпринимаются попытки кластеризовать партии и ходы.

Ключевые слова: компьютерный движок, кластеризация, рейтинг Эло

Для цитирования: Бобровская О. П., Лысенкова С. А. Математическое моделирование и анализ для определения силы игрока сёги // Вестник кибернетики. 2022. № 3 (47). С. 14–24. DOI 10.34822/1999-7604-2022-3-14-24.

Original article

MATHEMATICAL MODELING AND ANALYSIS TO DETERMINE A SHOGI PLAYER'S SKILL LEVEL

Olga P. Bobrovskaya^{1✉}, Svetlana A. Lysenkova²

^{1,2} Surgut State University, Surgut, Russia

¹ o-bobrovskaya@mail.ru ✉, <https://orcid.org/0000-0001-7045-9085>

² lysenkova_sa@surgu.ru, <https://orcid.org/0000-0003-1007-7610>

Abstract. The study deals with obtaining numerical estimates of shogi pieces' disposition on the board with the help of software. Using these estimates, the authors intend to evaluate a player's skill. That requires collecting data, reviewing and repeating existing approaches, proposing and implementing the authors' approach, and comparing the results. The difference in strength of two players is calculated based on their win distribution via re-simulation of the Ferreira approach. In implementing Yamashita's approach, the relation between Elo rating and the average degradation of the position estimate as a result of a bad move is searched for. The authors' approach is proposed to estimate a player's skill. Hence, the dependence of wins, good and bad moves, the deviation of the estimates of real and ideal moves, and the average improvement of the position from the Elo rating are analyzed. The study attempts to cluster plays and moves.

Keywords: computer engine, clustering, Elo rating

For citation: Bobrovskaya O. P., Lysenkova S. A. Mathematical Modeling and Analysis to Determine a Shogi Player's Skill Level // Proceedings in Cybernetics. 2022. No. 3 (47). P. 14–24. DOI 10.34822/1999-7604-2022-3-14-24.

ВВЕДЕНИЕ

Люди играют в шахматы (различного вида и с разными правилами) более двух тысячелетий. Компьютерные программы, называемые «движками» и способные играть в шахматы, создаются на протяжении полувека. Начиная с 70-х гг. прошлого века, Международная шахматная федерация осуществляет оценку силы игроков, составление их рейтинга и присвоение званий на основе результатов партий с использованием формул Эло. В случае победы рейтинг игрока увеличивается, в случае поражения – уменьшается на величину, зависящую от исходных значений рейтингов противников. Чем больше значение рейтинга игрока – тем он сильнее, чем больше разность рейтингов между двумя игроками – тем больше вероятность того, что сильнееший выиграет партию. Та же рейтинговая система используется в играх шахматного типа (в том числе сёги), также ее модификации используются во многих других играх.

Из-за недостатков, присущих рейтингу Эло, как в предсказании исходов партий, так и в скорости изменения, исследователями предпринимаются попытки найти лучшие альтернативы. Например, в работе [1] авторы разрабатывают нейронную сеть, способную предсказывать результат шахматной партии, основываясь на уже совершенных ходах (достигается точность предсказания 59 % для длины последовательности в 80 ходов).

С 2006 г. развивается идея оценки силы игрока по ходам партии с помощью компьютерных движков в шахматах: используются оценки движка для получения численной оценки позиции [2–5], количества возможных ходов (интерпретирующихся как сложность позиции) [2], определения соответствующих рейтингу Эло параметров чувствительности и согласованности [3], создания марковского классификатора, предсказывающего исход игры [4], подсчета распределения выигрыша (изменение в оценке позиции в результате хода) [5]. После чего полученные значения используются для создания альтернативы рейтингу Эло.

Согласно [6], в сёги до 2014 г. не существовало сравнений игроков на основе анализа игровых записей. Но автор начинает эту

работу и обнаруживает линейную зависимость между средним плохим ходом (изменением оценки позиции после хода игрока в худшую сторону) и рейтингом Эло. В статье [7] предлагается оценивать игроков через два параметра: значение среднего плохого хода (разница между оценкой хода и оценкой наилучшего хода от движка, если она больше 700 очков) и коэффициент совпадения (количество раз, когда лучший ход совпадает с фактическим ходом, деленное на количество ходов в партии), которые, как выяснилось, обязательно коррелируют между собой.

Цель данного исследования – проверить применимость уже существующих подходов и предложить свой подход определения силы игрока на основании его ходов в течение партии. Выдвигается гипотеза о том, что если ходы охарактеризовать численной оценкой улучшения или ухудшения позиции (что можно осуществить с помощью эталона – компьютерного движка), которые они производят, то все ходы, совершаемые в партиях, можно будет разбить на отдельные классы.

Для достижения цели следует выполнить следующие задачи:

1. Собрать и подготовить данные (записи партий).
2. Рассмотреть существующие подходы и разработать свою модель, в соответствии с этим обработать данные движком.
3. Повторить существующие подходы и реализовать свою модель.
4. Сравнить полученные результаты.

МАТЕРИАЛЫ И МЕТОДЫ

Объект исследования – партии сёги. Игра с нулевой суммой, с двумя игроками, которые ходят по очереди. После захвата фигуры противника, в свой ход ее можно выставить на поле как свою. В среднем в партии около 120 ходов, но в используемом наборе данных присутствуют как короткие игры (когда игрок сдался), так и длинные (вплоть до 1000 ходов).

В процессе исследования применялись теоретические методы: анализ и синтез, сравнение, абстрагирование и конкретизация, обобщение, формализация, моделирование; математические: статистические методы, методы визуализации данных (функции

и графики), методы искусственного интеллекта (машинное обучение).

Для численной оценки позиций использовался движок YaneuraOu с функцией оценки Elmo на глубине поиска 17. Все ходы партии отправлялись движку, чтобы получить от него 3 оценки: численную оценку получаемой в результате совершенного хода позиции, потенциально наилучший ход и численную оценку получаемой в его результате позиции. Выдаваемые движком значения оценок нормированы (не превышают по модулю 10000 единиц). Из пошаговой разницы полученных оценок позиций находились как оценки наилучших возможных ходов, так и реально совершенных.

Из собранного набора данных (с сайта 81dojo.com игры за 2011–2012 гг.) были выбраны пары игроков, у которых между собой больше 50 игр. Получилось 27 игроков. Для них, на основании имеющихся игр, посчитан процент выигрышей из всех сыгранных ими игр и проведен подсчет рейтинга Эло (начальное количество баллов 1500, как в работах [6, 7]). Полученный рейтинг Эло оказался слишком далек от рейтингов игроков, рассчитанных онлайн-платформой на основании значительно большего количества их игр. Поэтому для сравнений использовался текущий рейтинг Эло с сайта, несмотря на возможные изменения силы игры за 10 лет.

На основании полученных от компьютерного движка оценок позиций для каждого игрока подсчитан процент хороших, увеличивающих выигрыш, и плохих, ухудшающих позицию ходов, а также среднее количество очков, на которые изменяется оценка позиции в результате хороших и плохих ходов. Также подсчитано среднее значение разницы между оценкой наилучшего возможного хода, предлагаемого движком, и ходом игрока.

Данные, полученные с помощью компьютерного движка, используются для осуществления кластеризации методом k-средних. Для полученных кластеров оценивается сбалансированность. Для определения количе-

ства кластеров используются метод локтя, метод силуэта и Calinski-Harabasz index.

Подход Феррейры (Ferreira [4]). Расчетный рейтинг, вычисляемый Феррейрой, нельзя назвать точным рейтингом (каким является рейтинг Эло), поэтому автор интерпретирует его как силу игрока.

Для реализации подхода, описанного в статье [4], были использованы результаты игр двадцати семи пар игроков, у которых между собой больше 50 игр, и их рейтинг Эло. Для каждого игрока взята его последняя игра (предположительно, являющаяся наиболее показательной и соответствующей имеющемуся рейтингу Эло). На рис. 1 показаны гистограммы выигрыша за ход (разность между оценкой позиции до и после хода) для двух игроков. Значения частот на этих гистограммах были нормализованы, что означает, что сумма высот всех столбцов равна 1,0.

Приведенные на графиках значения интерпретируются как распределение выигрыша.

Далее, для игроков считается корреляция их распределений выигрыша по следующей формуле:

$$f_{X-Y}[n] = (f_Y \times f_X) \triangleq \sum_{m=-\infty}^{\infty} f_Y[m] \times f_X[n+m], \quad (1)$$

где X и Y – две дискретные случайные величины с распределениями вероятностей f_X и f_Y соответственно; тогда распределение вероятности $X - Y$ (победы первого игрока) определяется взаимной корреляцией.

Определяется вероятностная оценка выигрыша партии первым игроком:

$$p_{XY} \triangleq 0,5 \times P(X = Y) + 1 \times P(X > Y) = 0,5 \times f_{X-Y}[0] + \sum_{n=1}^{\infty} f_{X-Y}[n]. \quad (2)$$

Разница в рейтинговых баллах, соответствующая процентному ожиданию p , определяется по формуле:

$$d = \Phi^{-1}(p) \times 200 \times \sqrt{2}, \quad (3)$$

где Φ^{-1} – обратная функция плотности нормального распределения.

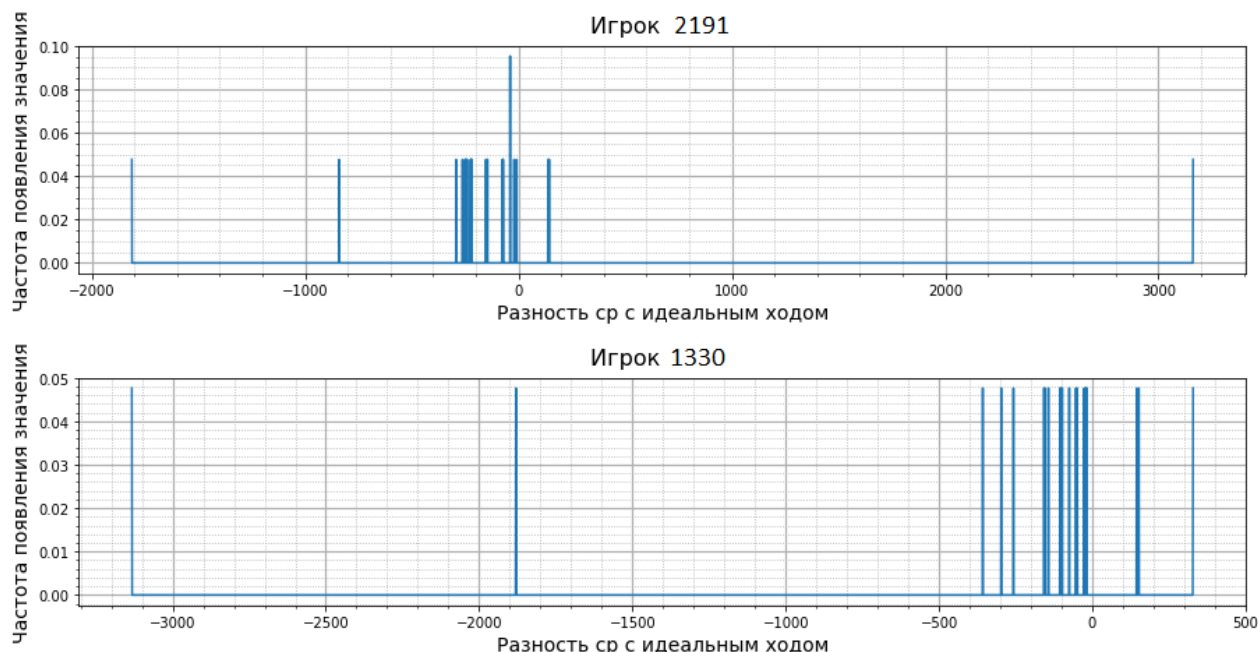


Рис. 1. Нормализованная гистограмма выигрыша для обоих игроков (строка № 23 из табл. 1)

Примечание: составлено авторами.

Такие расчеты были проведены для всех 27 игроков, которые могли оказаться как первым, так и вторым игроком в рассматриваемом наборе игр. Результаты представлены в табл. 1.

Таблица 1

Значения для крайних игр каждого рассматриваемого игрока

№	Индекс первого игрока	Индекс второго игрока	Рассчитанная разность рейтинга	Разность текущих рейтингов Эло
1	1	159	36,42	-33,6
2	1379	7	18,71	-264,53
3	9	407	-20,16	182,01
4	66	29	25,81	-490,1
5	66	187	51,92	-17,9
6	1330	204	98,6	92,9
7	1379	304	45,89	-238,7
8	9	407	-20,16	182,01
9	1330	780	67,09	-189,5
10	1330	789	1,18	-50,9
11	822	799	43,91	121,9
12	1330	884	27,33	-475,1
13	1379	1012	-33,58	255,03
14	1198	1377	-103,34	-651,8
15	1379	1230	26,14	-198,13
16	2191	1329	-36,72	807,45
17	2032	1377	92,44	-1153,42
18	2032	1379	-40,56	-322,42
19	1379	1432	146,63	-72,48
20	1438	1329	-44,61	635,57
21	1785	1379	32,65	16,3
22	2032	1379	-40,56	-322,42
23	2191	1330	-37,89	147,45

Примечание: составлено авторами.

Как видно из табл. 1, значения разности рейтингов Эло и рассчитанных разностей (по Феррейре) совпадают примерно в двух случаях из двадцати. Чтобы убедиться в нали-

чий или отсутствии связи между ними, был построен график по всем играм рассматриваемых игроков (рис. 2).

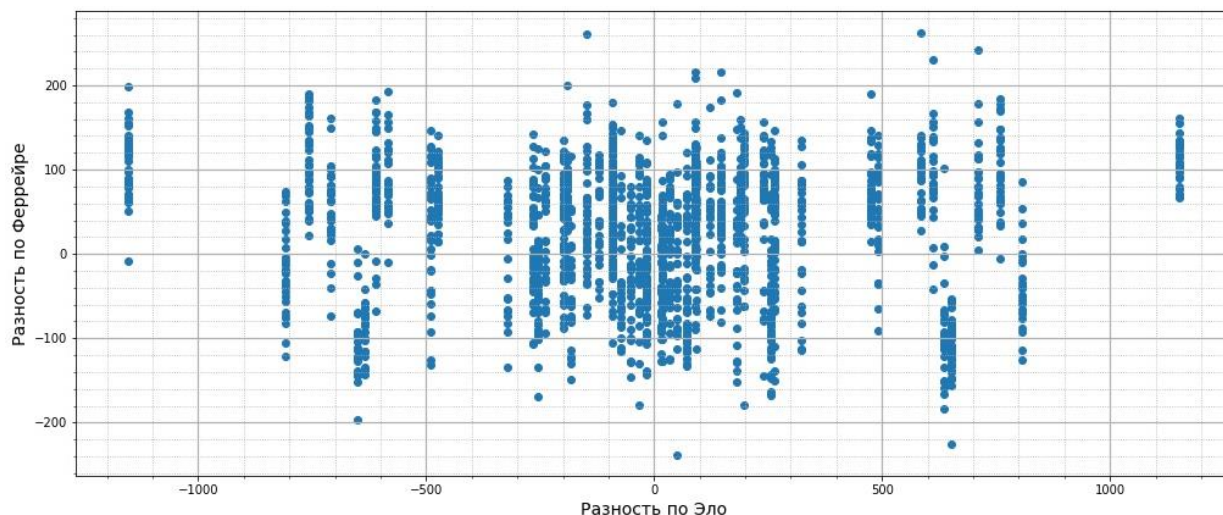


Рис. 2. Значения разности по Эло и по Феррейре для всех игр рассматриваемых игроков
Примечание: составлено авторами.

Значения по оси X фиксированы для игроков, а значения по оси Y рассчитываются для каждой партии, и, как видно по графику, сгруппированы в диапазоне около 200 единиц. Но как порядок величин, так и знак сильно отличаются: в то время как разница рейтинга Эло у игроков превышает 1000 ед. по модулю, разница по Феррейре не выходит за границы 300 ед. по модулю; значения симметричны относительно оси X как для положи-

тельной разности рейтинга Эло, так и для отрицательной, хотя можно было бы ожидать, что график будет сгруппирован в первой и третьей четвертях.

Подход Ямаситы. Попробуем повторить внешний вид графика (рис. 3) из статьи [6], на котором по оси X откладываются значения рейтинга Эло, а по оси Y – средний плохой ход.

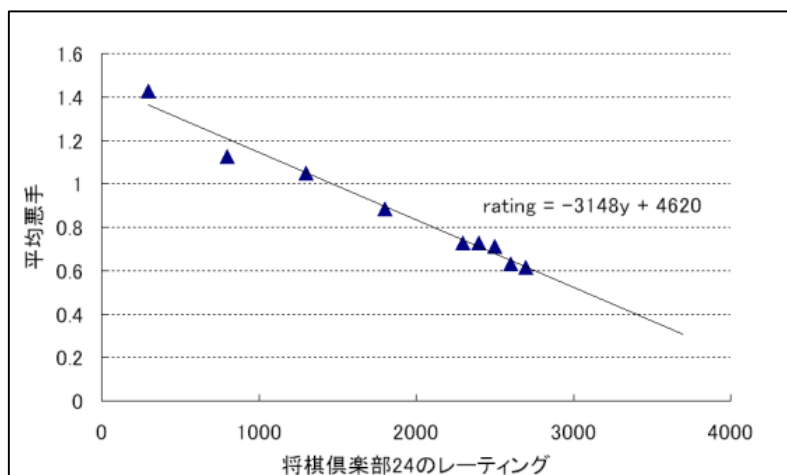


Рис. 3. Линейная зависимость значения рейтинга Эло от среднего плохого хода, полученная методом наименьших квадратов

Примечание: составлено по [6].

Значения рейтинга Эло были отсортированы по возрастанию для 27 игроков с количеством игр между собой больше 50 шт. Слева направо, сверху вниз на рис. 4 изображены графики: процент побед от количества сыгранных игр (рис. 4а); процент хороших ходов, в результате которых увеличивалась оценка позиции (рис. 4б); процент плохих ходов, в результате которых уменьшалась оценка

позиции (рис. 4в); средняя разница между реальным ходом и лучшим рекомендуемым ходом от движка (рис. 4г); среднее увеличение оценки позиции в результате хорошего хода (рис. 4д); среднее уменьшение оценки позиции в результате плохого хода (рис. 4е). Похожим на убывающую прямую должен быть шестой график.

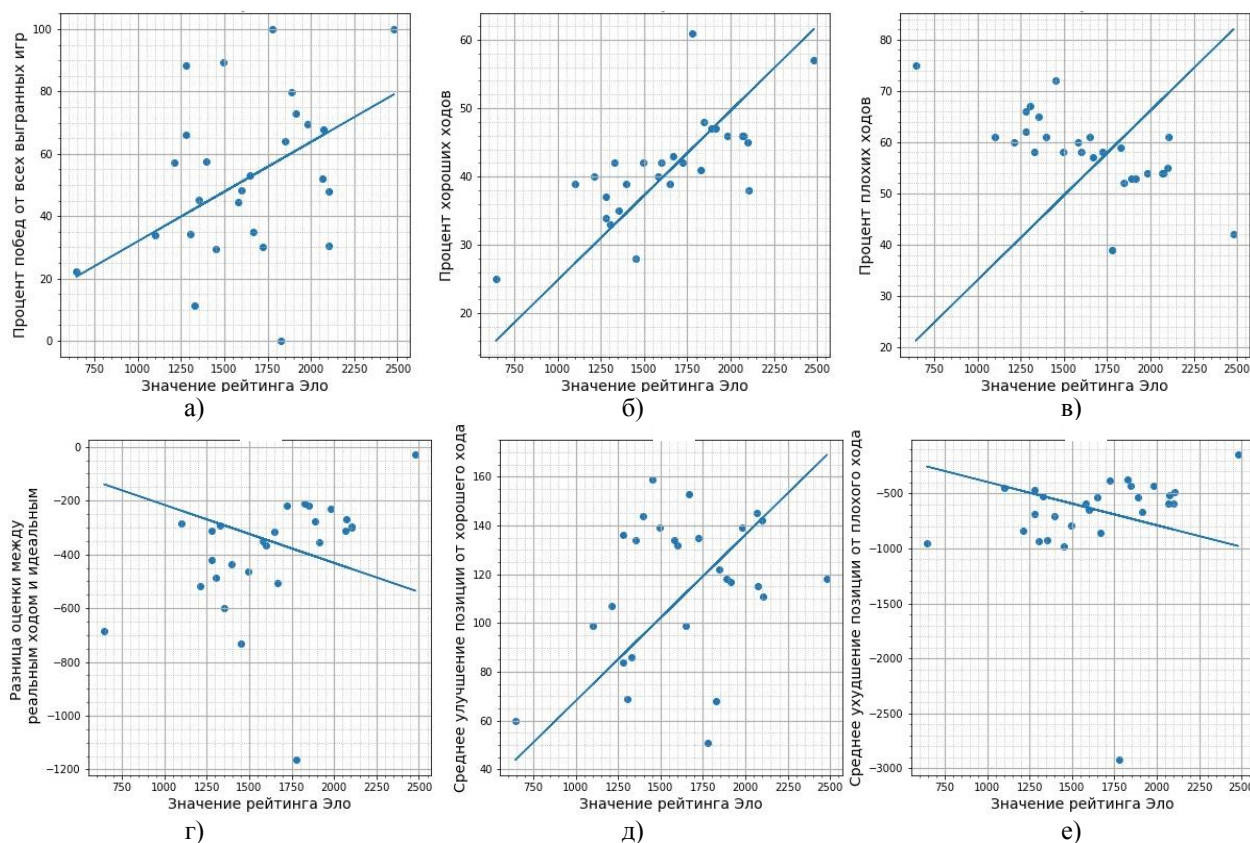


Рис. 4. Сравнение оценок игроков и их рейтинга Эло

Примечание: составлено авторами.

Прямые построены методом линейной регрессии. Для каждого графика был посчитан коэффициент R^2 (коэффициент детерминации). Для графиков под буквами: а, г, д, е, — значение коэффициента меньше 0,20, что говорит о большой разбросанности значений. Для графиков процентов хороших и плохих ходов (рис. 4б, 4в) значение коэффициента равно 0,53 и 0,54 соответственно, и, как видно по самому графику, точки лежат довольно близко к построенной прямой.

Кластеризация партий. Здесь используется только оценка позиции в результате

реального хода, на основе которой вычисляется ее изменение (разность последовательных значений). Последовательность значений изменения позиции для каждой партии представляет собой временной ряд. Для получения параметров этих рядов и последующей кластеризации был использован экстрактор признаков [8], обеспечивающий систематическое извлечение признаков временных рядов (комбинирующий установленные алгоритмы из статистики, анализа временных рядов, обработки сигналов и нелинейной динамики с надежным алгоритмом

выбора признаков). После разбиения на 4 кластера методом k-средних получены кла-

Таблица 2

Кластеризация партий

№ кластера	Количество партий
2	23846
0	9960
1	4219
3	1708

Примечание: составлено авторами.

Количество партий в кластерах значительно различается (на порядок), что, возможно, свидетельствует о принадлежности всех рассматриваемых партий к одному кластеру или крайне несбалансированных данных. Но, так как тут рассматривались все имеющиеся в нашем распоряжении партии (около 47 тысяч), можно предположить, что на основе данной кластеризации осуществлять последующую классификацию партий будет невозможно.

Так как в партии участвуют два игрока, было выдвинуто предположение, что удастся лучше разделить на кластеры не целые игры,

а половины игр, за которые отвечает каждый игрок. Игра разделена пополам – для каждого игрока его ходы: из последовательных разниц оценок позиции в одной партии строится два временных ряда: для первого и второго игроков (изменения численной оценки преимущества за 1-й, 3-й, 5-й и т. д. ходы; 2-й, 4-й, 6-й и далее ходы соответственно). Результат кластеризации представлен в табл. 3.

Результат стал чуть лучше, но различия в размерах кластеров все еще велики. Это позволяет утверждать, что кластеризовать партии не удалось.

Таблица 3

Кластеризация половин партий

№ кластера	Количество партий
3	40356
1	17431
2	12085
0	9594

Примечание: составлено авторами.

Кластеризация ходов. Кластеризация ходов осуществлялась по следующим параметрам: изменению оценки позиции, совпадению хода с наилучшим ходом от движка и разнице между ними.

В случае деления на 4 кластера (табл. 4) размеры кластеров получились сбалансированными (все числа одного порядка).

Таблица 4

Кластеризация ходов

№ кластера	Количество ходов
3	900244
1	889880
0	836716
2	783252

Примечание: составлено авторами.

Так как измерений немного, график с параметрами ходов представлен на рис. 5, где каждая точка – это ход. Темнота точек (от черного до светло-оранжевого) отражает номер хода в партии. По оси X отражено изменение оценки позиции, произошедшее в результате хода – выигрыш или проигрыш очков. По оси Y –

упущенный выигрыш (или проигрыш), который был бы возможен в случае выбора игроком наилучшего хода. Таким образом, точки, находящиеся под прямой $y = x$, являются ходами, в которых игрок сходил лучше рекомендуемого движком хода.

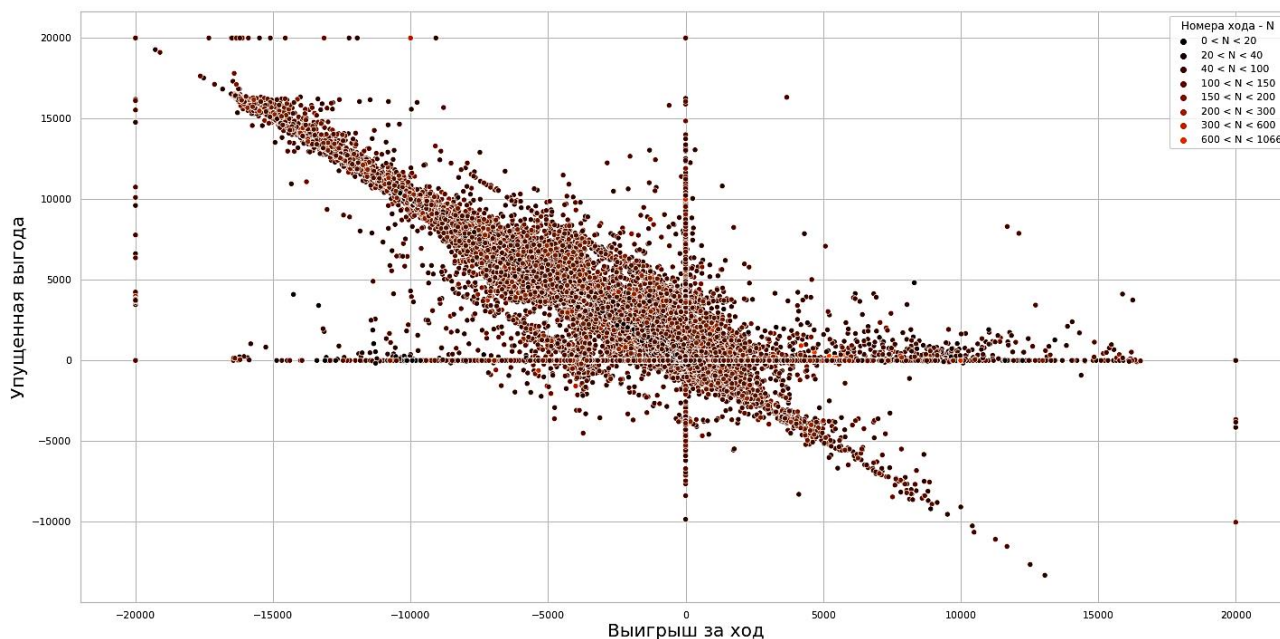


Рис. 5. Ходы

Примечание: составлено авторами.

Большая часть ходов находится выше прямой $y = x$ и в отрицательных значениях x , что говорит о том, что игроки преимущественно ходят хуже, чем могли бы, и вместо возможного в результате хода улучшения позиции ухудшают ее.

Для определения наилучшего количества кластеров были использованы метод локтя, метод силуэта и быстрее считающийся Calinski-Harabasz index, представленные ниже на графиках (рис. 6).

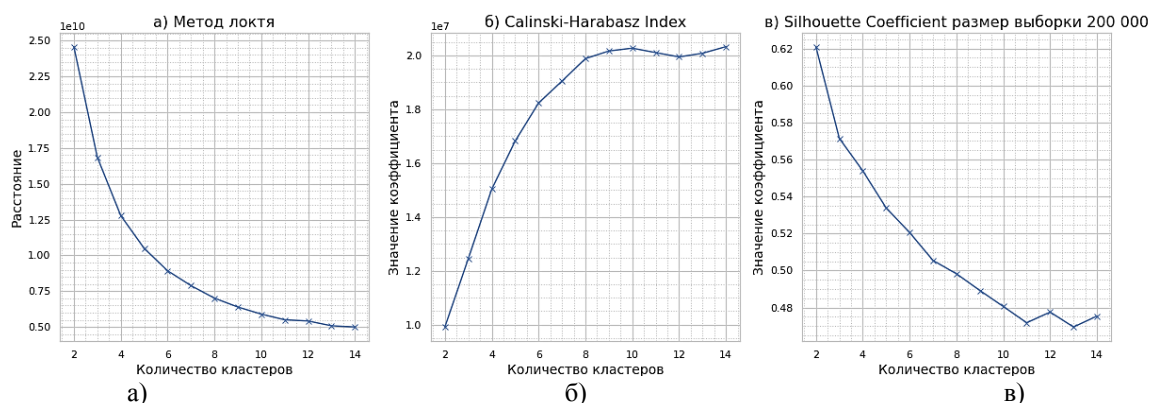


Рис. 6. Определение количества кластеров:

а – по методу локтя; б – Calinski-Harabasz index; в – по методу силуэта

Примечание: составлено авторами.

Метод локтя (Elbow rule) – визуальный метод определения наилучшего количества кластеров путем сравнения разницы суммы квадратов ошибок (SSE) каждого кластера; самая большая разница, образующая угол изгиба, показывает лучший номер кластера [9].

Calinski-Harabasz index известен как критерий отношения дисперсии. Индекс *Калински – Харабаз* определяется как отношение сумм межкластерной дисперсии и внутрикластерной дисперсии [10]:

$$CH = \frac{(T - W_K) / (K - 1)}{W_K / (N - K)}, \quad (4)$$

где $T = \sum_{i=1}^N \sum_{v \in V} (y_{iv} - \bar{y}_v)^2$ – разброс данных;

$W_K = W(S, C) = \sum_{k=1}^K \sum_{i \in S_k} d(y_i, c_k)$ – минимизируемое внутрикластерное расстояние между экземплярами кластера и центроидами. Наилучшим количеством кластеров считается то, у которого значение индекса наибольшее.

Метод силуэта [11] – метод для оценки потенциально оптимального числа кластеров, который использует среднее расстояние между одной точкой данных и другими в том же кластере и среднее расстояние между различными кластерами для оценки результата кластеризации. Метрика оценки этого метода называется коэффициентом силуэта (S) и определяется как $\frac{(b-a)}{\max(a, b)}$,

где a и b представляют собой среднее расстояние внутри кластера и среднее расстояние до ближайшего кластера соответственно. Лучшее значение равно 1, а худшее значение равно –1. Значения около 0 указывают на перекрывающиеся кластеры. Отрицательные значения обычно указывают на то, что образец был отнесен к неправильному кластеру, так как другой кластер более похож.

Метод локтя говорит о том, что кластеров от четырех до шести (плавное место перегиба графика), в то время как Calinski-Harabasz index и коэффициент силуэта не имеют четкого максимума, из чего можно сделать вывод, что произвести деление на кластеры методом k -средних в данном случае не удастся.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

В процессе исследования было обнаружено, что попытка рассчитать рейтинг Эло на малом количестве игр для малого количества игроков приводит к неадекватным результатам. Что подтверждает указанный вначале недостаток рейтинга Эло, требующего большого количества игр, и актуальность исследования, нацеленного на поиск альтернативы.

Полученные результаты не удается легко сопоставить с рейтингом Эло, на который мы опираемся в данном исследовании.

Подход Феррейры. Наблюдаемое как в табл. 1, так и на рис. 2 несоответствие знаков и масштабов разниц рейтингов Эло и Феррейры может иметь следующие возможные объяснения:

- неточности со стороны компьютерного движка, поскольку от запуска к запуску оценка одного и того же хода может меняться, чего, возможно, не наблюдалось у движка, используемого Феррейрой;

- неучтенные особенности шахмат и сёги: разность в средней продолжительности партии, характер этапов партии (дебют, миттельшпиль, эндшпиль) и переходов между ними, которые могут оказывать влияние на оценку рассматриваемых ходов;

- неточность рейтинга Эло как по временным промежуткам, так и по качеству для игроков-любителей;

- отмеченная автором [5] необходимость учитывать сложность позиций, возникающих за доской в ходе игры: «В конце концов, в некоторых позициях допустимых ходов всего несколько и ошибиться практически невозможно, а в других позициях сделать правильный выбор из множества возможных ходов становится очень сложно»;

- результаты в работе Феррейры по рассмотренной оценке разности рейтингов также не совпадали с настоящими рейтингами Эло игроков. Возможно, следовало продолжить анализ до получения финальных значений рейтинга по Феррейре, но отказ от этого мотивирован возможностью не использовать в анализе гипотезу о распределении движка с частотой 1 на значении выигрыша 0 (по наблюдениям, движок иногда выбирает ходы

не с самой большой оценкой) и последующем смещении результатов относительно известных рейтингов Эло.

Подход Ямаситы. В качестве положительного результата можно отметить, что точки значений на рис. 4б и рис. 4в расположены довольно близко к построенным прямым. Поэтому стоит рассмотреть идею использования процента хороших и плохих ходов, оцененных с помощью компьютерного движка, в качестве оценки силы игроков сёги.

Несоответствие графика на рис. 4е оригинальному графику (рис. 3) может быть объяснено неточностью повторения оригинального подхода, а именно:

- не были удалены первые 40 ходов построения крепости, не были убраны ходы, оценка которых превышала некоторый порог, чтобы исключить влияние форсированного мата (цуме);

- не исключены ходы, сумма абсолютных различий оценок которых при каждом переходе к лучшему ходу в ходе итерационного углубления превышают 1000 ед.

Кластеризация партий. Попытка кластеризовать партии как временные ряды не удалась. Предположение о том, что партии могут составлять классы, например по разности рейтингов противников или сочетанию стилей их игры, не нашло подтверждения. Если подобные классы и присутствуют в рассматриваемых данных, то в слишком малом количестве.

Кластеризация ходов. Как видно по графикам на рис. 5 и 6, деление на кластеры методом k-средних для рассматриваемых данных невозможно. Но ходы в отмеченных на графике (рис. 5) координатах можно разделить на области с помощью прямых, параллельных прямой $y = -x$. Полученные классы будут характеризоваться разной успешностью хода в сравнении с наилучшей оценкой движка.

Чтобы выделить из совершаемых игроками ходов различные классы методом k-средних, потребуется добавление большего количества параметров, характеризующих ход (промежуточные оценки, проводимые движком: статическая оценка позиции, все перерасчеты, проводимые в процессе поиска в глубину). Эти параметры можно либо достать из существующего движка с открытым кодом, либо разработать свой движок так, чтобы он выдавал эти оценки.

ЗАКЛЮЧЕНИЕ

Были подготовлены данные в виде записей партий, которые обрабатывались компьютерным движком для получения численных оценок позиций фигур на доске.

Рассмотрено несколько существующих подходов для определения силы игрока на основе их ходов, два подхода (Феррейры и Ямаситы) повторены. Достигнуть тех же результатов, что и у авторов подходов, не удалось.

Предложенные авторские подходы к оценке силы игрока (использование процента побед, разницы оценки между реальным и идеальным ходами и среднего улучшения позиции от хорошего хода, а также кластеризация партий и их половин) не дали удовлетворительных результатов.

Использование же процента хороших и плохих ходов игрока и кластеризация ходов являются перспективными. Поставленная в начале исследования цель – проверить применимость существующих и предложить свой подход для определения силы игрока на основании его ходов с помощью компьютерного движка – была достигнута.

Планируется продолжать работу над определением классов ходов, возможно, сопоставлением набора таких классов и рейтинга Эло игрока.

Список источников

1. Drezewski R., Wątor G. Chess as Sequential Data in a Chess Match Outcome Prediction Using Deep Learning with Various Chessboard Representations // Procedia Computer Science. 2021. Vol. 192. P. 1760–1769.

References

1. Drezewski R., Wątor G. Chess as Sequential Data in a Chess Match Outcome Prediction Using Deep Learning with Various Chessboard Representations // Procedia Computer Science. 2021. Vol. 192. P. 1760–1769.

2. Guid M., Bratko I. Computer Analysis of World Chess Champions // ICGA Journal. 2006. Vol. 29, No. 2. P. 65–73.
3. Regan K. W., Haworth G. M. Intrinsic Chess Ratings // Proceedings of the AAAI Conference on Artificial Intelligence. 2011. Vol. 25, No. 1. P. 834–839.
4. Alliot J.-M. Who is the Master? // ICGA Journal. 2017. Vol. 39, No. 1. P. 3–43.
5. Ferreira D. R. Determining the Strength of Chess Players Based on Actual Play // ICGA Journal. 2012. Vol. 35, No. 1. P. 3–19.
6. Yamashita H. Pro Shogi Players Rating and Game Record Analysis // Proceedings of the Game Programming Workshop. 2014. P. 9–16.
7. Takatsu K., Takada H., Hirata T. How to Estimate the Rating of the Youngest Professional Shogi Player Souta Fujii // Memoir of Graduate School of Engineering, University of Fukui. 2019. Vol. 67. P. 1–9.
8. Christ M., Braun N., Neuffer J., Kempa-Liehr A. W. Time Series Feature Extraction on Basis of Scalable Hypothesis Tests (Tsfresh – A Python Package) // Neurocomputing. 2018. Vol. 307. P. 72–77. URL: <https://www.sciencedirect.com/science/article/pii/S0925231218304843> (дата обращения: 23.05.2022).
9. Umargono E., Suseno J. E., Gunawan S. K. V. K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula // Proceedings of the 2nd International Seminar on Science and Technology (ISSTEC 2019). 2020. P. 121–129. URL: https://www.researchgate.net/publication/346349075_K-Means_Clustering_Optimization_Using_the_Elbow_Method_and_Early_Centroid_Determination_Based_on_Mean_and_Median_Formula (дата обращения: 29.05.2022).
10. De Amorim R. C., Hennig C. Recovering the Number of Clusters in Data Sets with Noise Features Using Feature Rescaling Factors // Information Sciences. 2015. Vol. 324. P. 126–145. URL: <https://arxiv.org/pdf/1602.06989.pdf> (дата обращения: 29.05.2022).
11. Shi C., Wei B., Wei S. et al. A Quantitative Discriminant Method of Elbow Point for the Optimal Number of Clusters in Clustering Algorithm // J Wireless Com Network. 2021. P. 31. DOI 10.1186/s13638-021-01910-w (дата обращения: 29.05.2022).
2. Guid M., Bratko I. Computer Analysis of World Chess Champions // ICGA Journal. 2006. Vol. 29, No. 2. P. 65–73.
3. Regan K. W., Haworth G. M. Intrinsic Chess Ratings // Proceedings of the AAAI Conference on Artificial Intelligence. 2011. Vol. 25, No. 1. P. 834–839.
4. Alliot J.-M. Who is the Master? // ICGA Journal. 2017. Vol. 39, No. 1. P. 3–43.
5. Ferreira D. R. Determining the Strength of Chess Players Based on Actual Play // ICGA Journal. 2012. Vol. 35, No. 1. P. 3–19.
6. Yamashita H. Pro Shogi Players Rating and Game Record Analysis // Proceedings of the Game Programming Workshop. 2014. P. 9–16. (In Japanese).
7. Takatsu K., Takada H., Hirata T. How to Estimate the Rating of the Youngest Professional Shogi Player Souta Fujii // Memoir of Graduate School of Engineering, University of Fukui. 2019. Vol. 67. P. 1–9. (In Japanese).
8. Christ M., Braun N., Neuffer J., Kempa-Liehr A. W. Time Series Feature Extraction on Basis of Scalable Hypothesis Tests (Tsfresh – A Python Package) // Neurocomputing. 2018. Vol. 307. P. 72–77. URL: <https://www.sciencedirect.com/science/article/pii/S0925231218304843> (accessed: 23.05.2022).
9. Umargono E., Suseno J. E., Gunawan S. K. V. K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula // Proceedings of the 2nd International Seminar on Science and Technology (ISSTEC 2019). 2020. P. 121–129. URL: https://www.researchgate.net/publication/346349075_K-Means_Clustering_Optimization_Using_the_Elbow_Method_and_Early_Centroid_Determination_Based_on_Mean_and_Median_Formula (accessed: 29.05.2022).
10. De Amorim R. C., Hennig C. Recovering the Number of Clusters in Data Sets with Noise Features Using Feature Rescaling Factors // Information Sciences. 2015. Vol. 324. P. 126–145. URL: <https://arxiv.org/pdf/1602.06989.pdf> (accessed: 29.05.2022).
11. Shi C., Wei B., Wei S. et al. A Quantitative Discriminant Method of Elbow Point for the Optimal Number of Clusters in Clustering Algorithm // J Wireless Com Network. 2021. P. 31. DOI 10.1186/s13638-021-01910-w (accessed: 29.05.2022).

Информация об авторах

О. П. Бобровская – магистрант.

С. А. Лысенкова – кандидат физико-математических наук.

Information about the authors

O. P. Bobrovskaya – Master's Degree Student.

S. A. Lysenkova – Candidate of Sciences (Physics and Mathematics).