Strategies and first-absorption times in the random walk game

M. I. Krivonosov[1,2]✉, *S. N. Tikhomirov*[1]

[1]National Research Lobachevsky State University of Nizhny Novgorod, Russia
[2]Ivannikov Institute for System Programming of the RAS, Moscow, Russia
E-mail: ✉krivonosov@itmm.unn.ru, stikhomirov@itmm.unn.ru
Received 22.10.2022, accepted 5.04.2023, available online 15.05.2023, published 31.05.2023

Abstract. *Purpose* of this work is to determine the average time to reach the boundaries, as well as to identify the strategy in the game between two players, controlling point movements on the finite square lattice using an independent choice of strategies. One player wants to survive, i. e., to stay within the interior of the square, as long as possible, while his opponent wants to reach the absorbing boundary. A game starts from the center of the square and every next movement of the point is determined by independent strategy choices made by the players. The value of the game is the survival time that is the number of steps before the absorption happens. In addition we present series of experiments involving both human players and an autonomous agent (bot) and analysis of the survival time probability distributions. *Methods.* In this work, methods of the theory of absorbing Markov chains were used to analyze strategies and absorption times, as well as the Monte Carlo method to simulate trajectories. Additionally, a large-scale field experiment was conducted using the developed mobile application. *Results.* The players' strategies are experimentally obtained for the cases of playing against an autonomous agent (bot), as well as human players against each other. A comparison with optimal strategies and a random walk is made: the difference between the experimental strategies and the optimal ones is shown, however, the resulting strategies show a much better result of games than a simple random walk. In addition, especially long-running games do not show the Markovian property in case of the simulation corresponding strategies. *Conclusion.* The sampled histograms indicate that the game-driven walks are more complex than a random walk on a finite lattice but it can be reproduced with a Markov Chain model.

Keywords: random walk, markov chain, random walk game, mobile application, game experiment.

Стратегии и время поглощения в игровых случайных блужданиях

М. И. Кривоносов[1,2] ✉, *С. Н. Тихомиров*[1]

[1]Национальный исследовательский
Нижегородский государственный универсистет им. Н. И. Лобачевского, Россия
[2]Институт системного программирования им. В. П. Иванникова РАН, Москва, Россия
E-mail: ✉krivonosov@itmm.unn.rum, stikhomirov@itmm.unn.ru

Аннотация. *Цель* настоящего исследования — определить среднее время достижения границы, а также выявить стратегии в игре между двумя игроками, управляющими движением фишки на конечной квадратной решетке с помощью независимого выбора стратегий. Один игрок старается оставаться внутри квадрата как можно дольше, пока его противник старается достичь поглощающей границы. Игра начинается в центре квадрата, и каждое следующее движение фишки определяется стратегиями, выбираемыми игроками независимо друг от друга. Результат игры — это время выживания, то есть количество шагов до того, как произойдет поглощение. Дополнительно в работе представляются результаты проведения серии экспериментов с участием как игроков-людей, так и автономного агента (бота), и анализ соответствующих распределений вероятностей времени выживания. *Методы.* В данной работе применялись методы теории поглощающих марковских цепей для анализа стратегий и времен достижения границы, а также метод Монте-Карло для симуляции траекторий. Дополнительно были применены подходы к проведению масштабного полевого эксперимента с использованием разработанного мобильного приложения. *Результаты.* Экспериментально получены стратегии игроков для случаев игры против автономного агента (бота), а также игроков-людей друг против друга. Проведено сравнение с оптимальными стратегиями и случайным блужданием, в ходе которого показано отличие экспериментальных стратегий от оптимальных, однако полученные стратегии показывают значительно лучший результат игр, чем простое случайное блуждание. Дополнительно проанализированы особенно длительные игры, не обладающие свойством марковости при столкновении соответствующих стратегий. *Заключение.* Найденные распределения указывают на то, что исследуемый процесс является более сложным, чем случайное блуждание на конечной решетке, однако распределение может быть воспроизведено с помощью моделей цепи Маркова.

Ключевые слова: случайные блуждания, марковские цепи, игровые случайные блуждания, мобильные приложения, игровой эксперимент.

Introduction

A one-dimensional random walk on a finite discrete interval is famously interpreted as a story of two gamblers playing a coin-flipping game again and again until one of the gambler's capital is exhausted and he is ruined [1, 2]. In the random-walk framework, this is a situation when the random walker — whose position is determined by the current capital of one of the players — reaches one of the two boundary points and is absorbed there. The ruin time is therefore first-passage or survival time, one of the key concepts of the random walk theory [3].

Assume now that the game is not completely random but the corresponding transition probabilities are controlled by the players. This is the idea of "games of survival" proposed by Hausner [4] and Peisakoff [5] in their notes in 1952. Similar to the original gambler ruin's process [6], survival games can be generalized to two and three dimensions and finally formulated in the most general way as a game-driven motion over some finite domain with a border [7]. In such a game one of the players tries

Кривоносов М. И., Тихомиров С. Н.
Известия вузов. ПНД, 2023, т. 31, № 3

335

to avoid hitting the border while his opponent wants to reach the border as fast as possible. The game is specified with a matrix with a certain number of rows (strategies of the first player) and columns (strategies of the second player). On every step, the players independently choose their strategies, that is the first player chooses a row and the second one — the column, and the move of the point is determined by the corresponding element of the matrix. The game is over once the point has reached the boundary.

Even though the idea of such games was formulated more then seventy years ago [7], no quantitative results were presented until now. This is not a surprise as such games are characterized by high complexity [8]. Here we present an experimental realization of a two-dimensional survival game. Our idea is to consider the game as a first-passage process [3] and analyze the collected game histories from the corresponding perspective, by inspecting histograms of the survival time.

From the very start, random walks were interpreted in biological and behavioral terms, and now Pearson's drunkard [9] is no less famous than Schrödinger's cat. Different types of random walks [10, 11] are used as foundations to build optimal search theories and explain the spatial-temporal patterns produced by foraging animals and people exploring the Disney World park [12].

However, whenever a new random-walk model is introduced, a round of discussion about model's validity and the importance of all details neglected during the model's formulation immediately starts. E.g., it is often mentioned that search landscapes are not uniform but patchy and often exhibit fractal-like patterns, which might induce a power-law scaling in foraging trajectories [13, 14], or that people usually follow streets or existing trails which form a very correlated structure [15], etc. In other words, there are always factors and features that are not taken into account and thus there is always room for a discussion on whether these features are important and relevant; see, f.e., Refs. [16, 17]. Therefore, it would be beneficial to *design* a process, perform series of experiments (with organisms or humans), and collect the statistically sounding amount of data — which then can be analyzed.

In case of humans, game theory provides a possibility to implement this idea. The concept of game-driven or "game-type" [7] random walks, when players determine the move of a point by independently choosing strategies, looks like a good opportunity.

Here we present the results collected for a two-dimensional game-driven process, by using a mobile application (app) "Random Walk Game" [18]. The app provides two people with a possibility to play regardless of their location, interrupt games, and challenge each other for a game by inviting the opponent from the list of registered players. The app allowed us to collect data more quickly and with less effort than traditional methods (such as lab sessions) usually demand. Additionally, the app provides a human player with a possibility to play anytime against a bot (the latter, so to say, is simply flipping a fair coin on every round).

We analyze the collected game histories from a point of view of survival time statistics, by comparing the histograms of the game duration with the distributions corresponding to random walks and distributions obtained with Markov-chain models. The experimental histograms bear several features which demonstrate that the game-driven walks can be, in general, emulated with a Markov-chain model. However, we also have detected some games with anomalous length whose occurrence cannot be explained as 'rare events' of the corresponding model.

1. Game

The Random Walk Game is a game played on a square lattice by two players, A and B. The lattice consists of vertices (x, y), $x, y \in \{-N/2, ..., -1, 0, 1, ..., N/2\}$, where N is an even number, and the size of field is $n \times n$, where $n = N + 1$. The game always starts from the center of the square, i.e., from the vertex $(0, 0)$. Both players can see the current position of the marker on the lattice. The boundary vertices of the lattice can be defined as absorbing points and the game terminates once the marker gets into one of these vertices.

336

Кривоносов М. И., Тихомиров С. Н.
Известия вузов. ПНД, 2023, т. 31, № 3

Players control the movement of the marker by making an independent choice of one out of the two possible strategies. Information about the possible movement is represented by a matrix (analog of a pay-off matrix in the standard games) which is known by both players; see Fig. 1, *a*. The movement of the marker is defined by the mutual choice of strategies; Fig. 1, *b*, Fig. 2.
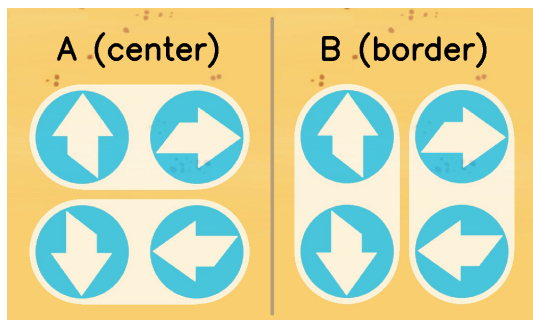
a
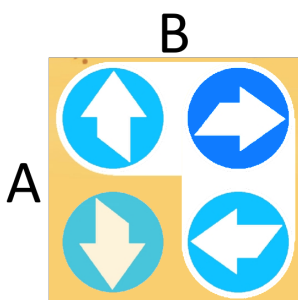
b

Fig. 1. *a* — The strategies of two players, player A (playing 'for the center') and player B (playing 'for the border'). Each player has two strategies to choose from: Player A — (up, right) and (down, left) and player B — (up, down) and (right, left). *b* — Example of the one turn. At each step of the game, players independently choose their strategies. For example, if player A has chosen the first button (up, right) and the player B has chosen the second button (right, left), the next move will be to the right (color online)
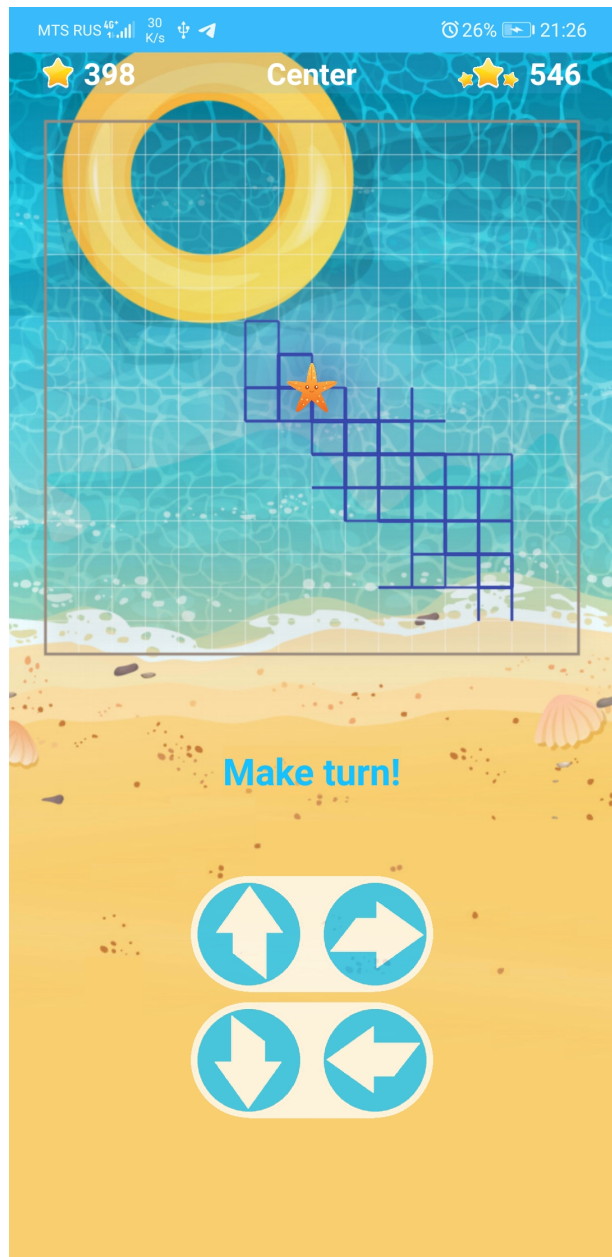
Fig. 2. Screenshot of the Random Walk Game. The header shows of the number of turns made (left), what the player is playing for (in this case they are 'playing the center', i.e., they try to avoid the absorption for as long as possible) (center), and the number of turns in the longest up-to-date game (right). Players see the game field, the marker current position and the trajectory of the game. At the bottom of the screen there are the strategies for a player to choose from. In this case, these are the rows of the movement 'pay-off' matrix (color online)

Кривоносов М. И., Тихомиров С. Н.
Известия вузов. ПНД, 2023, т. 31, № 3

337

The aim of player A is to keep the marker inside the square, i.e., at one of the internal vertices, and thus avoid the absorption as long as possible. We say that this player 'plays the center'. The aim of player B is to reach the absorbing boundary, i.e., one of the boundary vertices, as fast as possible. We say that this player 'plays the border'.

A typical trajectory produced during a game is shown on Fig. 2.

2. Experiment

2.1. Mobile application.   To carry out the experiment, we have developed a mobile application [18] which offers two game types: (i) to play against another player (P vs P) or (ii) to play against the bot (P vs B, in case the player plays for the center or B vs P, in case the player plays for the border). The bot makes a completely random choice of its strategy at every turn. The app transmits player choices over the Internet to the web-server and receives back the resulting movements. All player choices and resulting movements are collected in the server database for further analysis.

Choices of the bot in P vs B and B vs P games are determined by the sequence of pseudo-random numbers computed by using Mersenne Twister random number generator (implemented in PHP 7.4 as mt_rand function). The game is played on a square lattice 17×17 (which means $N = 16, n = 17$).

2.2. Players.   Players, that are the participants of the experiment, are aged from 16 to 52 years. Players in the category <25 years are volunteers from the students of Lobachevsky State University and Higher School of Economics. Players above 25 are volunteers from academic staff of different universities and research institutions in Russia, Germany, France, Norway, and South Korea.

2.3. Experiment.   Player versus player (P vs P) sessions were organized in the three following formats.

1. Students were gathered in the same room and played without communications for one hour. The participant's goal was to obtain the highest score in the room. Player pairs were formed based on similarity of players' skills in the competitive programming.
2. Students were gathered in the same room and played without communications for two hours. The participant were arranged by the sum of weighted scores of played games. The longer game — the higher weight associated with this game for player A (and the opposite for player B).
3. Players played without communications for around 30 minutes per day, whenever they wanted to play. The aim is to reach the highest cumulative score.

To organise Player vs Bot game sessions we used two formats.

1. Participants played against the bot at will (they were also able to pause the game any moment). This was continued for one month. Players were then ranked by the ratio of exponentially moving average (EMA) score of playing 'for the center' (P vs B game) to the EMA score of playing 'for the border' (B vs P game). Player ranking table was then made available online to all participants.
2. Participants played against the bot at will. The ranking table defined as the result of the implementation of the previous format was used. The task for a participant was to beat the highest score.

Whether a player is supposed to stay inside the square as long as possible (i.e., Player A type) or reach the absorbing boundary as soon as possible (Player B type), is decided at random at the beginning of every game. Altogether, we have collected a database consisting of 512 realizations for the P vs P game and 528 and 534 realizations for the P vs B and B vs P games, respectively.

338

Кривоносов М. И., Тихомиров С. Н.
Известия вузов. ПНД, 2023, т. 31, № 3

3. Analysis

We are interested in the statistics of the game duration, i.e., the number of turns k before the absorption on the boundary happens and the game terminates. In terms of the random walk theory this will correspond to the 'first-passage time' or 'survival time' [3].

We will consider three different types of games: (i) a pure random game, when two players choose their strategy at random (B vs B game or random walk), (ii) a game when a player plays against the environment (P vs B and B vs P games) and (iii) a game played by two players (P vs P game).

3.1. Absorbing Markov Chains. To analyze P vs B, B vs P, and P vs P games, we use the framework of Markov Chain theory [19].

The state of the system is a coordinate vector $w_k = (x_k, y_k)$ at the k-th turn, where position corresponds to the vertex and coordinates vary in the ranges: $-N/2 \leqslant x_k \leqslant N/2$, $-N/2 \leqslant y_k \leqslant N/2$. Overall there are $n^2 - 4$ reachable vertices (excluding 4 unreachable corners), $r = 4(n-2)$ of which correspond to absorbing states ($|x_k| = N/2 \vee |y_k| = N/2$) and the remaining $s = (n-2)^2$ to transient states ($|x_k| < N/2 \wedge |y_k| < N/2$), see Fig. 2. Although, there are two parity cases of the lattice size, we consider only odd sizes due to the symmetry of the lattice with respect to the center. We arrange vertices into a column according to the horizontal scanning of the lattice. After that we introduce transition matrix [20] $P((i,j) \rightarrow (x,y))$ with elements corresponding to probabilities to go from the state (i,j) to (x,y) after one turn.

The absorbing Markov Chain theory provides a way to compute survival time in the case of absorption processes [20]. According to the theory, transition matrix P can be represented as a block matrix:

$$P = \begin{bmatrix} Q & R \\ \mathbf{0} & I_r \end{bmatrix} \tag{1}$$

where Q is an s-by-s matrix corresponding to transitions between internal states, R is a non-zero s-by-r matrix corresponding to transitions from internal states to absorbing states, and last two matrices are corresponding to loops in the absorption states: $\mathbf{0}$ is an r-by-s zero matrix, and I_r is the r-by-r identity matrix [20].

Here we resort to the standard concept of the mean absorption time which is based on the idea of fundamental matrix [20] (it also appears in the context of quasistationary state theory [21]).

Namely, we construct fundamental matrix L:

$$L = \sum_{k=0}^{\infty} Q^k = (I_s - Q)^{-1}, \tag{2}$$

where I_s is the s-by-s identity matrix. The (i,j) entry of matrix L is the expected number of times the chain is found in state j, given that the chain started in state i.

Based on the property of fundamental matrix, we obtain the expected number of steps before being absorbed starting in transient state i:

$$T = L\mathbf{1}, \tag{3}$$

where $\mathbf{1}$ is a column vector whose entries are all 1.

Since the starting position of the game is a center of square lattice the resulting expected number of steps before being absorbed t_n corresponds to entry of T for the $(0,0)$ state, where n is a size of field.

3.2. Game as a Markov Chain. To model the game as a Markov chain, we have to specify transition probabilities between different states. In general, relying on memoryless property of Markov chains, we define mixed strategy S_{ij}^p determined by Bernoulli distribution σ_{ij}^p over two pure strategies for the player $p \in \{A, B\}$ in the state (i, j):

$$\sigma_{ij}^p(s) = \begin{cases} f_{ij}^p, & \text{if } s = 0, \\ 1 - f_{ij}^p, & \text{if } s = 1, \end{cases} \tag{4}$$

where $f_{ij}^p \in [0, 1]$ is a probability of picking first strategy $(s = 0)$.

This in turn allows us to determine the transient probability from the state (i, j) to (x, y) for the corresponding mixed strategies:

$$P\left((i, j) \to (x, y)\right) = \begin{cases} 0, & \text{if } |i - x| + |j - y| \neq 1, \\ f_{ij}^A\left(1 - f_{ij}^B\right), & \text{if } x = i \wedge y = j + 1, \rightarrow \\ \left(1 - f_{ij}^A\right) f_{ij}^B, & \text{if } x = i + 1 \wedge y = j, \downarrow \\ \left(1 - f_{ij}^A\right)\left(1 - f_{ij}^B\right), & \text{if } x = i \wedge y = j - 1, \leftarrow \\ f_{ij}^A f_{ij}^B, & \text{if } x = i - 1 \wedge y = j, \uparrow \end{cases} \tag{5}$$

The connectivity in the system is defined by transitions to the four neighboring states on the square lattice. This way, the first case of transient probability denotes the zero probability between unconnected states, and others are transition probabilities to neighboring states in the corresponding direction.

The case of the random strategy choice corresponds to equal probability to choose strategies for both players, $f_{ij}^A = 1/2$, $f_{ij}^B = 1/2$. Therefore, the transient probability is equal to $1/4$ for neighbour state transition and zero otherwise. Having applied the absorbing Markov chain formalism, we were able to compute the mean absorption time t_n^{BvsB} depending on the field size n.

The last two cases, games against the bot (P vs B and B vs P games) and games of two players (P vs P games) explicitly determine only the bot strategy which is the equiprobable choice. In contrast, strategies of participants are unknown. Therefore, in order to assess behavior we assume the absence of memory in participant choices and estimate relative frequencies of their strategies depending on the current position of the marker on the lattice. Next, we interpreted the experimentally obtained relative frequencies as probabilities and computed the mean absorption time for both cases, t_{17}^{PvsB}, t_{17}^{BvsP} and t_{17}^{PvsP}.

3.3. Probability evolution model. Although absorbing Markov Chains let us find the mean duration of the game, the exact distribution of the duration cannot be obtained in this framework. To solve this problem, we define a mathematical model based on a Markov chain property of the game.

Let us first define probability W_{ij}^k of finding a marker in the state (i, j) on the k-th turn. The $(0, 0)$ is the initial state of the marker and for $k = 0$ the probability to find the marker in this position $W_{00}^0 = 1$ and for all other states $W_{ij}^0 = 0$, $(i, j) \neq (0, 0)$. By sequentially applying modified transition matrix $\widetilde{\text{P}}$ to state probability vector $\mathbf{w^k}$, we could propagate the latter:

$$\mathbf{w^{k+1}} = \widetilde{\text{P}}\mathbf{w^k}, \quad k \geqslant 0 \tag{6}$$

340

Кривоносов М. И., Тихомиров С. Н.
Известия вузов. ПНД, 2023, т. 31, № 3

where $\widetilde{\mathrm{P}}$ is n^2-by-n^2 transition matrix P modified in such a way that reaching the boundary state excludes the marker from the system. These modification are required to compute exact probability W_{ij}^k that marker finishes in the particular absorption state at turn k, where $(i,j) \in \mathbf{B}$ — boundary states.

Then, the probability of marker being absorbed on the k-th turn can be computed using the following expression:

$$p_{\text{abs}}^k = \sum_{(i,j)\in\mathbf{B}} W_{ij}^k, \tag{7}$$

where $\mathbf{B}$ is a set of boundary states.

The analysis described above was applied to all three cases, B vs B, P vs B (also B vs P), and P vs P.

3.4. Numerical simulations. To model game realizations, we perform stochastic simulations of the game process by using transition matrices constructed from the collected database.

The mark's start is located at the state $(0,0)$ initially and then, at each turn, its move is determined by sampling of strategy choices, Eq. (4). Upon selecting the strategies, the marker is moved according to Eq. (5). Eventually, marker reaches an absorption vertex and the simulation stops.

We compute statistical properties and distributions considered in this Section in order to compare the obtained results with the modelling and experimentally obtained statistics. All three approaches (numerical simulations, probability evolution model, and inversion of fundamental matrix) and their statistics were implemented using Python 3.8 and performed on a work station (Core i5-8600 3.1GHz, 32Gb RAM).

The source code is available at https://github.com/SermanVS/RWAnalyzer.

3.5. B vs B game or random walk. This game results in the standard random walk on finite two-dimensional lattice, initiated at the central vertex of the lattice. There is the analytic closed form solution for the mean absorption time [6]:

$$\overline{k}_N^{\text{BvsB}} = \frac{2}{N} \sum_{k=0}^{N/2-1} (-1)^k \alpha_{k,N} \frac{\cos((2k+1)\pi/(2N))}{\sin^3(2k+1)\pi/2N},$$
$$\alpha_{k,N} = 1 - \frac{1}{\cosh\left(N/2 \cosh^{-1}\left(1 + 2\sin^2\frac{(2k+1)\pi}{2N}\right)\right)}. \tag{8}$$

For our case, $N = 16$, this expression gives 75.22. Additionally, we simulated the process to sample the probability distribution of k. It is shown by the purple line on Fig. 4, *a*, *b* (Cf. below).

4. Results

Aiming at analysis process, Python 3.8 is used for modelling, simulations and inversion of the matrices. Results address the 3 cases of the random walk game of 2 players on the finite lattice. Comparison to experimentally obtained trajectories of real games were done with using the 17×17 field. The size of the field was chosen due to the comfortable game time: longer game leads to fatigue and smaller games have simpler strategies. The players have spent approximately 250 hours in total playing in the Random Walk Game to produce experimental data. As a result, we have obtained trajectories of marker movements and corresponding player choices of 1562 real games in different modes (Cf. first row of the Table).

Table. The absorption mean times obtained by different approaches for the 3 game modes: B vs B is a pure random walk on 2D finite lattice, player versus bot in case of two goals: keep marker inside the field as long as possible (P vs B, center), reach the border as soon as possible (B vs P, border), and P vs P mode is a game of two players (all games and only games with length greater than 400 turns). The values are provided for the field size 17×17. The statistics in the simulation were performed using 10^5 trajectories. The modelling were performed up to 10^4 steps. The simulation, modelling and AMC theory (absorption Markov chain) were applied to the frequencies obtained from the real games in the experiment. In B vs B case were used the equiprobable choice strategy

Field size 17×17	Random walk	P vs B	B vs P	P vs P	P vs P 400+
# of games	–	528	534	500	13
Experiment	–	145.45	71.12	120.60	594.27
Simulation	75.22	145.77	73.66	115.93	132.91
AMC theory	75.21	145.85	73.79	116.22	133.22
Modelling	75.21	145.85	73.79	116.22	133.22
Optimal	–	225.00	64.00	N/A	–

4.1. Mean absorption time. The analysis of B vs B and P vs B (also B vs P) cases of random walk game suggests the quadratic function of the mean absorption time depending on the field size. The corresponding dependencies are presented on Fig. 3. In case of P vs B the identified relation for player A (goal: center) is $t_n^{\text{PvsB}} = (n-2)^2$ and for player B (goal: border) is $t_n^{\text{BvsP}} = (n-1)^2/4$. Though, relation for B vs B can be approximated by formula $t_n^{\text{BvsB}} = 0.294685413(n-1)^2 - 0.232$ with mean absolute error less than 10^{-3} at the range $[3; 1001]$. Coefficients at leading term n^2 are approximately equal to 1.0, 0.3, 0.25 in the order P vs B (goal center), B vs P (goal border), B vs B. Asymptotically the expected time of the game in the optimal center strategy (P vs B) is 3 times longer than the absorption time in the pure random walk. However, the optimal border strategy (B vs P) is only about 20% quicker than the pure random walk.

The mean absorption time in B vs B case was computed based on formula (8) is equal to 75.20846497681383 that match up to 14 significant digits to the values obtained by the inverse of fundamental matrix approach 75.20846497681377.
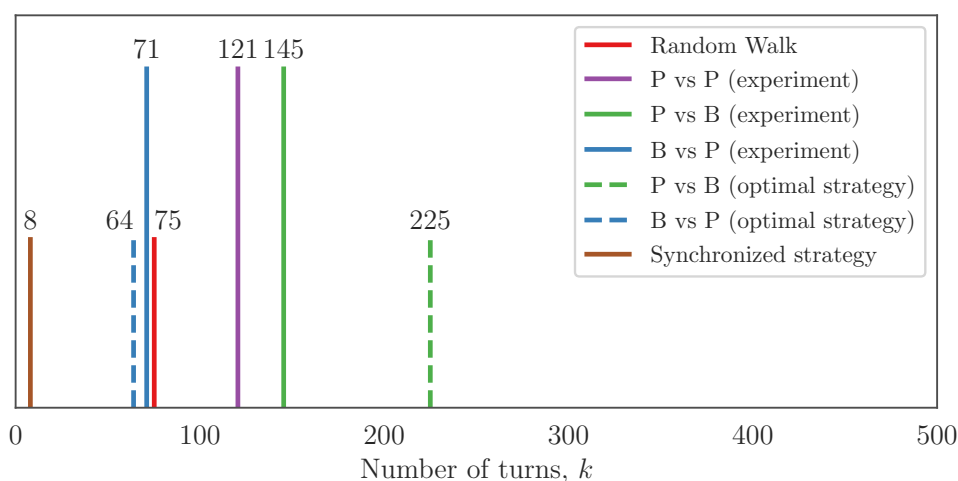

Fig. 3. Mean absorption times obtained in the experiment compared to different strategies and pure random walk (color online)

342

Кривоносов М. И., Тихомиров С. Н.
Известия вузов. ПНД, 2023, т. 31, № 3

Here, we showed the algorithm of the optimal strategy construction in B vs P case. It is evident that increasing the number of states on one-dimensional segment leads to higher mean absorption times. Therefore, in case of requirement to reach the border as soon as possible it is profitable to choose the smallest one-dimensional segment on the 2d plane. There are 2 shortest segments that passes through the center: vertical and horizontal. According to the rules of movements (Cf. Fig. 1), player B can explicitly choose one of those segments and keep the marker on this segment until reaching the border. The resulting process is a one-dimensional random walk on the shortest segment that has the property of shortest mean absorption time among other segments. The length of this segment is equal to n and based on the random walk theory the mean absorption time is equal to $t_n^{\mathrm{BvsP}} = (n-1)^2/4$. Considering of other 2 dimensional strategies leads to increased number of states on the path to the absorption states resulting in increased mean absorption time. The formal proof was done only for small sizes $5, 7$ by solving the global optimization problem using Wolfram Mathematica.

On the other hand, the goal of staying inside the field as long as possible requires the longest path for the random walk. The longest segment that can be placed within field corresponds to diagonal states. Although, such states are ambiguous, their constitute main and side diagonal "stairs". The movement rules enable player A to keep marker only on the main diagonal regardless of choices of the second player. That produces a random walk on the diagonal with length $2n-3$ because the states under the main diagonal have symmetry according to movement rules. Then, mean absorption time is equal to $t_n^{\mathrm{PvsB}} = ((2n-3-1)/2)^2 = (n-2)^2$.

Another strategy for player A can be described as follows: in neighbors of absorption states choose pure strategy with one probability corresponding to movement away from the field border, and in other states choose any strategy. Analysis of this algorithm was done for the sizes $5, 7, 9, 11$. As a result of the t_n simplification we found the independence of the mean absorption time from choices away from the border. The values of mean absorption times obtained by this algorithm matched to the optimal times of one-dimensional diagonal strategy t_n^{PvsB}. The formal proof of optimality of this algorithm will be interesting part of the future works.

Next step in the analysis of the expected time of game endings is the comparison of the theoretical optimum to the experimental statistics. The detailed experimental study description was provided in the section "Experiment". In total we obtained 1062 games in the player versus bot mode: 528 of which are corresponds to playing with a center goal (P vs B) and 534 — with a border goal (B vs P). The participants who should have kept marker inside the field as long as possible in the average shows 145.45 turns per one game. However, the optimal strategy suggests the mean absorption time equal to $t_{17}^{\mathrm{PvsB}} = (17-2)^2 = 225$, that 57% higher than obtained in the real games. Although, players showed lower average times than optimal, their games are still 2 times longer than pure random walk that provides 75.2 mean absorption time. The experiment shows the ability of the considered population to recognize properties of the game and to improve the simple strategy of random choices at each step regardless of knowledge about optimal strategies. However, the population strategy did not converge to the optimal strategy in the average. Applying the modelling approach to the average population strategy in the P vs B mode provides similar result of 145.85 absorption time.

Results of players aimed to reach border as soon as possible demonstrate the average number of turns equal to 71.12. This is 7 turns ahead compared to the optimal strategy $t_n^{\mathrm{BvsP}} = (17-1)^2/4 = 64$ and 4 turns faster than pure random walk. Also, in that game mode we see slight improvement of equiprobable random choice strategy and gap with the optimal strategy. Eventually, modelling of the average population strategy for B vs P border mode demonstrates quite higher mean absorption time 73.79 in contrast to the modelling of P vs B center strategy. This difference can be caused by inaccurate frequencies of rarely visited states.

Using frequencies sampled with the player versus bot games as input probabilities for the modelling Markov process which propagates probability distribution (Cf. Sections 3.2, 3.3, Eq. (5)) and

for the stochastic simulation procedure of individual trajectories (Cf. Section 3.4) we correctly reproduced the results. Therefore, frequencies of corresponding strategy choices determined for each state allowed us to interpret set of f_{ij} as the average strategy revealed by population. Simulation and modelling values are presented in the summary table of mean absorption times (Cf. the Table).

Finally, results of participant games with each other (P vs P) shows intermediate mean absorption times between P vs B center and B vs P border values. The average number of turns in P vs P games is 57 moves longer than mean absorption time in the optimal strategy of the B vs P mode with border goal. Comparison to the P vs B mode with center goal shows 25 steps lower values in P vs P versus experimental P vs B games and approximately 2 times lower values in P vs P versus optimal P vs B center strategy.

Mean absorption times obtained with simulations for particular frequencies coincide with AMC theory mean values (error is less than 10^{-9}). Therefore, both methods can be used to calculate mean absorption times for the arbitrary strategies. The simulation of 10^5 trajectories shows good error, accurate to the integer part.

4.2. Absorption time distributions. Both modelling and simulation approaches allowed us to compute precise and estimated probabilities of finishing the game at the arbitrary turn number. Applying these methods requires specifications of mixed strategy depends on the game field position. The pure random walk defines simple strategy of the equiprobable choice from pure strategies. Also, we analysed two suggested strategies of optimal random walk in the player versus bot mode. Besides the analytic strategies we collected experimental games of choice frequencies at each state. Based on those strategies we propagate probabilities through the field space and time. Additionally, we run simulations of 10^5 individual trajectories using different strategies. As a result we obtain distributions of absorption times that are presented on the Fig. 4.

All the distributions have differences in probabilities on even and odd turns on the close inspection. To assess only behavior associated with major trends, we consider histogram with wide bins (size 16). All the modes follow the same pattern of distribution:

1. the short games are rare;
2. distribution has one statistic mode at an intermediate number of turns;
3. the probability of long games exponentially decreases as game length increases.

The P vs B center mode demonstrates similar distribution form except the statistic mode corresponds to short games that remain hidden with chosen bin size.

Although the number of trajectories in the P vs P mode are small (500), very long games were observed in the distribution with length greater than 400 turns. The probability of obtaining such games according to modelling the population average strategy is lower than 0.015. However, 13 long trajectories were found in games of participants in the range from 461 to 964 turns and average absorption time equal to 594.27. Discovered abnormality can be explained by "synchronization" between individual brain activities in long runs. As one turn takes 4.5 seconds in the average, game with 400 turns is 30 minutes long. Such a long time of the failure to finish the game by player B causes frustration and reduces concentration that can lead to unconscious decisions that can be easily predicted by player A. A loss of concentration results in worse ability of the individual to produce pure random independent choices. In this regard, such "synchronization" events can arise in the game.

Analysis of these games was performed by decoupling strategies of the games from the main part of the distribution. Eventually, the modelling of the opposite average long run strategies produced the lower mean absorption time, approximately 133.22 turns. In order to compare distributions of such games we also modeled propagation of frequencies of movement directions that were observed in the long run games. The corresponding distributions of long run games are depicted in Fig. 5. Although,
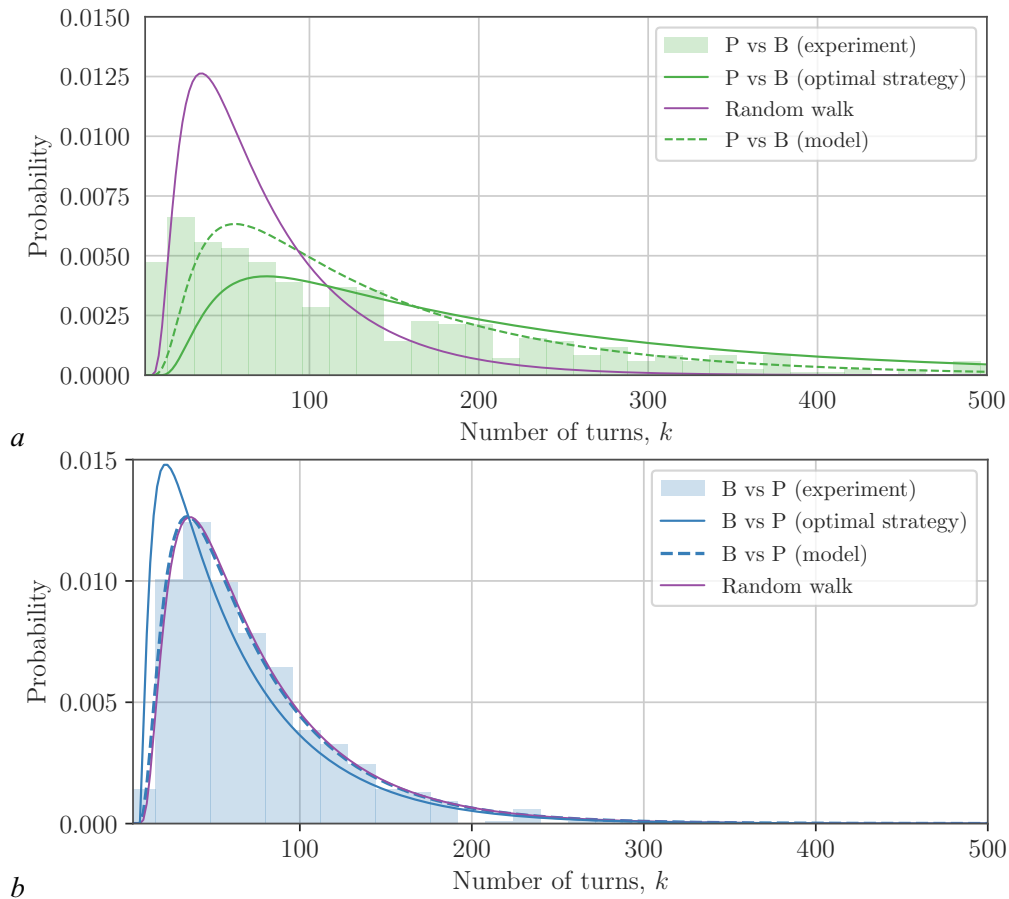
Fig. 4. Distributions of the number of turns obtained by modelling (solid line) and by simulation (dots) using corresponding strategies of players A and B. The Random walk mode (purple line) run equiprobable choice for both players. The player versus bot (green — P vs B, and blue — B vs P) curves were produced based on corresponding average population strategies. The bot chooses a strategy equiprobably. The optimal strategy for the P vs B center mode (green dashed line) is to keep a marker on the diagonal stair. The optimal strategy for the B vs P border mode (blue dashed line) is to only chooses movements along the horizontal line. Experimentally obtained histograms are presented for player versus bot modes (green area — P vs B, and blue area — B vs P) (color online)
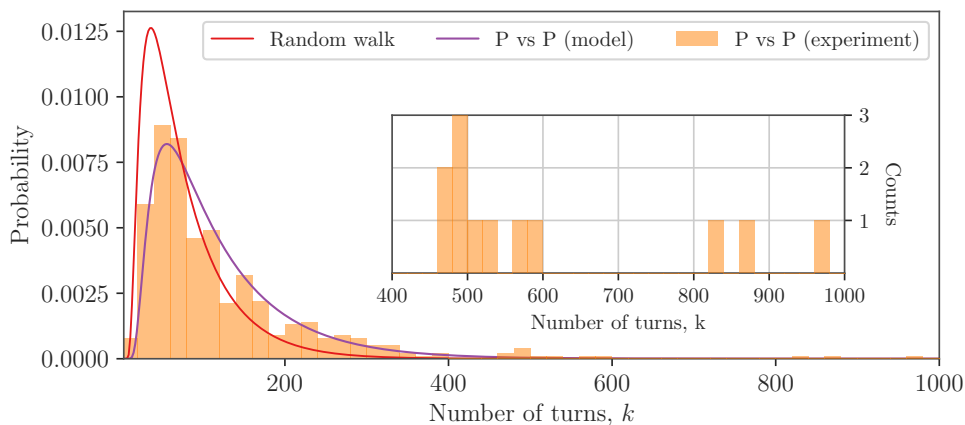

Fig. 5. The absorption time distributions for the P vs P mode (yellow histogram and purple line) compared to modelling of movement frequencies (green line) and strategies (blue line) observed in long run games (400+ turns). The frequencies of movement directions for each state obtained in the experimental long run games were used to propagate probabilities over the field by the modelling approach. Strategies of both players A and B in the P vs P with length over 400 turns were used separately in the modelling approach (color online)

Кривоносов М. И., Тихомиров С. Н.
Известия вузов. ПНД, 2023, т. 31, № 3

345

the distribution of frequencies demonstrates the long tail, the underlying strategies that appears in the "synchronized" games did not reproduce the emergence of long games. Therefore, it demonstrates the presence of choice dependence on hidden factors that can not be explained only by the Markov chain property.

4.3. Spatial distributions. Next, we analysed probabilities to find a marker in a particular state during the game. Similarly to previous sections we assessed experimentally obtained frequencies and results of modelling. Visualizations of spatial distributions for corresponding strategies are depicted on the Fig. 6.

The pure random walk demonstrates generalized multinomial distribution over space. In contrast, addition of game rules changes resulting distribution picture. Predictably, the P vs B mode with center goal demonstrates mostly movements on the diagonal (Fig. 6, *a*). Although, this behavior coincides with the spatial distribution in the suggested optimal strategy, the experimental data have higher spread around the diagonal states than optimal.

The optimal strategy in B vs P border mode should produce vertical or horizontal lines of involved states. As a result of experimental B vs P border games we obtained 3 main patterns in the distribution (Fig. 6, *b*): the expected movements on the straight lines, and different from the optimal pattern is a distribution similar to the pure random walk on 2D square lattice. The second pattern shows the attempt of population to find the best strategy in the two-dimensional space rather than in one-dimensional line. However, the leaving of one-dimensional lines increases the mean absorption times. This explains
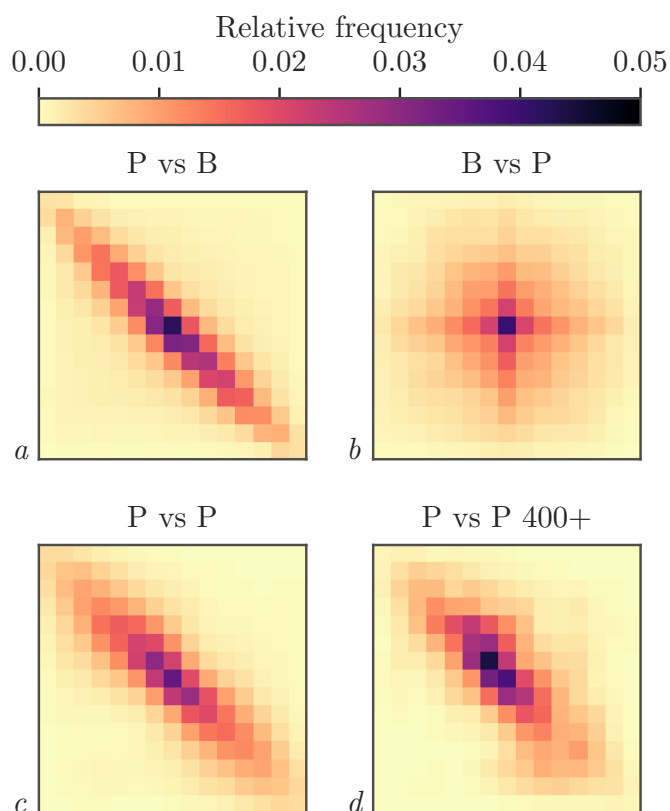

Fig. 6. The 2D distributions of state visits obtained in experimental games for 4 cases: *a* — P vs B with center goal, *b* — B vs P with the border goal, *c* — games of two players (P vs P), and *d* — games longer than 400 turns in P vs P mode (color online)

the small difference in terms of mean absorption times between experimental B vs P border and pure random walk (B vs B).

The spatial distribution pattern in the P vs P mode (Fig. 6, *c*) is similar to the P vs B center case: the marker mainly moves on the diagonal. The only difference is the enlarged spread around the diagonal line that shows stronger average strategy of player B with border goal compared to the equiprobable random choices (as in the P vs B center mode).

Lastly, we analysed the spatial distribution of states in games longer than 400 turns. In general, the probability of marker locations mostly concentrated on the main diagonal similar to previous cases (Fig. 6, *d*). Nevertheless, the distribution is more compact at the field center and has higher variation around the main diagonal compared to the P vs B center and P vs P cases. Such behavior suggests longer staying of marker close to the center during the long games, but also with movements from the diagonal and not only along it.

4.4. Strategy analysis.   Finally, we disentangle average population strategies of each player A and B and compare it to each other and to the optimal one. To present strategies we visualize them as the colored two-dimensional matrix with indices corresponding to the states on the 2D squared lattice (Fig. 7). Elements of the matrix depict the frequency of choosing the first pure choice from possible
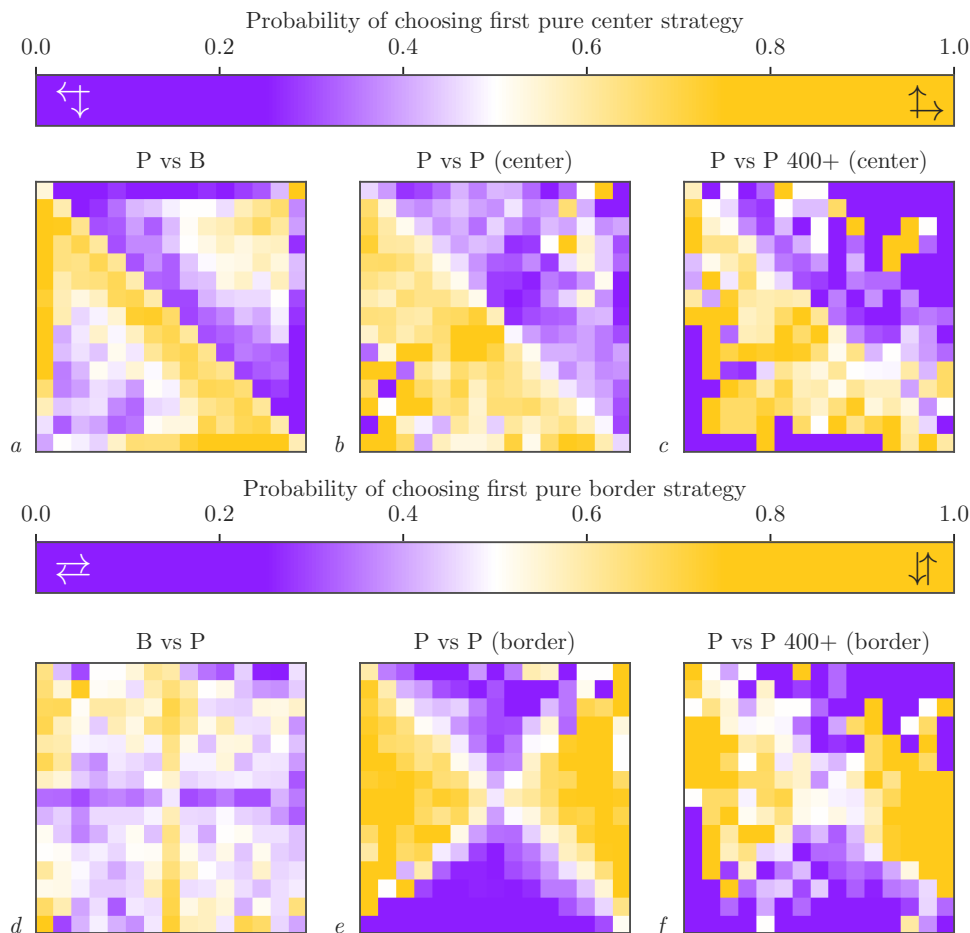

Fig. 7. The visualization of average population strategies for different modes obtained in the experiment. The color of cells depicts the frequency of choosing first pure strategy: for the center goal (*a*, *b*, *c*) and for the border goal (*d*, *e*, *f*) (color online)

choices according to the rules (Cf. Fig. 2). Technically, player A with center goal has two choices: move up/right, and move down/left; player B with border goal has another two choices: move up/down, and move left/right. The divergent colors of elements demonstrate which pure strategy dominates at each state in the average over all games.

The observed strategy in case of P vs B center mode shows mainly movement in direction of the diagonal states (Fig. 7, *a*). However, states distant from the diagonal shows slightly higher frequencies of the opposite choices. Generally, players choose go away from the border, but in the states closer to corners on the side diagonal the behavior becomes more random (relative frequencies are closer to 0.5). The experimental strategy on the main diagonal looks similar to the first optimal strategy. In contrast, the choices on the border differs from the second optimal strategy that suggests always moving away from the border. This results in a leak of probabilities outside the field not only in corners of main diagonal, but also in other border states.

The strategy of player with center goal in the P vs P case also shows similar strategy to the P vs B center mode (Fig. 7, *b*). Moreover, players on average choose to move in the direction of the main diagonal regardless of the position on the lattice. Almost all states at the border suggests similar strategy except some states with almost equal frequencies of both choices. In comparison to the P vs B center mode, player in P vs P mode with center goal have slightly lower confidence about moving towards main diagonal (the frequencies are closer to 0.5 in the P vs P center compared to P vs B center). As an example of experimental P vs P trajectory see Fig. 8.

Рис. 8. Trajectory of P vs P mode game obtained from the experiment divided by rounds that consist of 40 turns (color online)

Eventually, the B vs P border strategy is well-defined at the horizontal and vertical lines (Fig. 7, *d*). Although, players more often chose move towards the closest border at the central straight lines, frequencies in other states are not following to the common pattern. In contrast to the similarity of P vs B and P vs P center goal strategies, the strategy of P vs P border demonstrates very different picture with clear pattern. In this case players act exactly the opposite to the B vs P border mode. Their average strategy suggests the more often to choose moving along straight line that already conquered. That means movements along the coordinate line with lower deviation from the center. As a result plane of decisions is divided into 4 alternated triangles. Although, the frequencies are close to 0 or 1, they are not equal to. The small difference demonstrates rare attempts of movements towards the closest border.

Lastly, strategies obtained in the long games between 2 players did not produce clear patterns due to limited number of games. However, there is similarity of decision principles to the normal P vs P model.
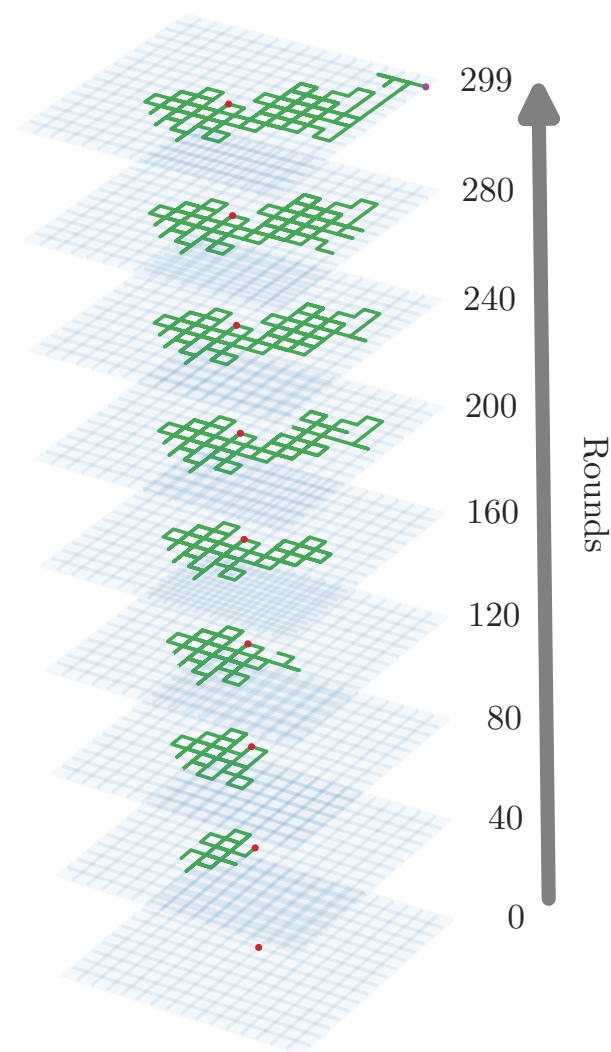
Conclusions

We presented the results of experimental studies of a game-driven spatial process which appears as the result of the interaction of two antagonistic players. The collected trajectories were analyzed as realizations of a two-dimensional random walk in a finite domain with absorbing boundaries. We chose to focus on the statistics of survival time and its features such as the shape of the corresponding probability distribution. We use a novel technique to conduct experiments and collect the data, based on a mobile app [22].

We demonstrated that the survival time histograms, obtained from the experimental data, can be reproduced with Markov Chain (MC) models. However, we also detected instances of super-long games which cannot be explained as 'rare events' of the Markov Chain models — because the probability to observe such a *single* event is negligibly small for the corresponding MC model. These 'anomalous' games were played by experienced players and their duration cannot be explained by, f.e., assuming that one of the players is a 'dummy' who chooses the next move (strategy) completely at random — in this case the duration of game can be estimated with a MC model and should be much shorter.

The origin of super-long games is a challenging open problem. We could only speculate that the mechanism behind is related to some sort of (anti)correlations appearing between antagonistic players and its explanation can only be found by going beyond the MC models and standard single-round theory of games.

As the next step, we are aimed at collecting more realizations thus obtaining more instances of super-long games. A detailed analysis of the corresponding trajectories can shed light on the origin of the counter-intuitive inter-player (anti)correlations.

References

1. Coolidge JL. The gambler's ruin. Annals of Mathematics. 1909;10(4):181–192. DOI: 10.2307/1967408.
2. Feller WD. An Introduction to Probability Theory and Its Applications. Wiley: New York; 1950. 704 p.
3. Redner S. A Guide to First-Passage Processes. Cambridge: Cambridge University Press; 2001. 312 p. DOI: 10.1017/CBO9780511606014.
4. Hausner M. Games of Survival. Report No. RM-776. Santa Monica: The RAND Corporation; 1952. 7 p.
5. Peisakoff MP. More on Games of Survival. Report No. RM-884. Santa Monica: The RAND Corporation; 1952. 20 p.
6. Kmet A, Petkovšek M. Gambler's ruin problem in several dimensions. Advances in Applied Mathematics. 2002;28(2):107–118. DOI: 10.1006/aama.2001.0769.
7. Romanovskii IV. Game-type random walks. Theory of Probability & Its Applications. 1961;6(4): 393–396. DOI: 10.1137/1106051.
8. Nisan N, Roughgarden T, Tardos E, Vazirani VV. Algorithmic Game Theory. Cambridge: Cambridge University Press; 2007. 754 p. DOI: 10.1017/CBO9780511800481.
9. Pearson K. The problem of the random walk. Nature. 1905;72:294. DOI: 10.1038/072294b0.
10. Zaburdaev V, Denisov S, Klafter J. Lévy walks. Reviews of Modern Physics. 2015;87(2):483–530. DOI: 10.1103/RevModPhys.87.483.
11. Bénichou O, Loverdo C, Moreau M, Voituriez R. Intermittent search strategies. Reviews of Modern Physics. 2011;83(1):81–129. DOI: 10.1103/RevModPhys.83.81.
12. Rhee I, Shin M, Hong S, Lee K, Kim SJ, Chong S. On the Levy-walk nature of human mobility. IEEE/ACM Transactions on Networking. 2011;19(3):630–643. DOI: 10.1109/TNET.2011.2120618.
13. Fauchald P. Foraging in a hierarchical patch system. The American Naturalist. 1999;153(6): 603–613. DOI: 10.1086/303203.

14. Scanlon TM, Caylor KK, Levin SA, Rodriguez-Iturbe I. Positive feedbacks promote power-law clustering of Kalahari vegetation. Nature. 2007;449(7159):209–212. DOI: 10.1038/nature06060.

15. Reynolds A, Ceccon E, Baldauf C, Karina Medeiros T, Miramontes O. Lévy foraging patterns of rural humans. PLOS ONE. 2018;13(6):e0199099. DOI: 10.1371/journal.pone.0199099.

16. Pyke GH. Understanding movements of organisms: it's time to abandon the Lévy foraging hypothesis. Methods in Ecology and Evolution. 2015;6(1):1–16. DOI: 10.1111/2041-210X.12298.

17. LaScala-Gruenewald DE, Mehta RS, Liu Y, Denny MW. Sensory perception plays a larger role in foraging efficiency than heavy-tailed movement strategies. Ecological Modelling. 2019;404:69–82. DOI: 10.1016/j.ecolmodel.2019.02.015.

18. Krivonosov MI, Tikhomirov SN. Random Walk Game [Electronic resource]. 2020. Available from: https://play.google.com/store/apps/details?id=com.scigames.RWGame (Google Play Store), https://apps.apple.com/us/app/random-walk/id1564589250 (AppStore).

19. Taylor HM, Karlin S. An Introduction to Stochastic Modeling. San Diego: Academic Press; 2008. 648 p.

20. Kemeny JG, Snell JL. Finite Markov Chains. New York: Springer-Verlag; 1983. 226 p.

21. Darroch JN, Seneta E. On quasi-stationary distributions in absorbing discrete-time finite Markov shains. Journal of Applied Probability. 1965;2(1):88–100. DOI: 10.2307/3211876.

22. Zhang J, Calabrese C, Ding J, Liu M, Zhang B. Advantages and challenges in using mobile apps for field experiments: A systematic review and a case study. Mobile Media & Communication. 2018;6(2):179–196. DOI: 10.1177/2050157917725550.

Кривоносов Михаил Игоревич — родился в Нижнем Новгороде (1995). Окончил с отличием Институт информационных технологий математики и механики Нижегородского государственного университета им. Н. И. Лобачевского по направлению «Прикладная математика и информатика» (2018). Окончил аспирантуру Института информационных технологий математики и механики по специальности «Математическое моделирование, численные методы и комплексы программ» (2022). С 2023 года работает в Институте системного программирования им. В. П. Иванникова Российской академии наук в должности младшего научного сотрудника и в центре фотоники ННГУ им. Н. И. Лобачевского в должности младшего научного сотрудника. Научные интересы — случайные блуждания, таксис, нейробиология, сложные сети, анализ временных рядов, компьютерные науки. Опубликовал свыше 30 научных статей по указанным направлениям.

603950 Нижний Новгород, пр. Гагарина, 23
Нижегородский государственный университет им. Н. И. Лобачевского
E-mail: krivonosov@itmm.unn.ru
ORCID: 0000-0002-1169-5149
AuthorID (eLibrary.Ru): 1066771

Тихомиров Сергей Николаевич — родился в г. Бор (1999). Выпускник Института информационных технологий, математики и механики Нижегородского государственного университета им. Н. И. Лобачевского по направлению «Математика и компьютерные науки» (2021). Обучается в магистратуре Института информационных технологий, математики и механики по специальности «Прикладная математика и информатика» (2023). С 2023 года работает в Центре фотоники ННГУ им. Н. И. Лобачевского в должности лаборанта-исследователя. Научные интересы — случайные блуждания, машинное обучение, анализ данных.

603950 Нижний Новгород, пр. Гагарина, 23
Нижегородский государственный университет им. Н. И. Лобачевского
E-mail: stikhomirov@itmm.unn.ru
ORCID: 0000-0001-8203-7090