Algorytmy ewolucyjne w Security Games

Adam Żychowski

Gry obronne Stackelberga

- Security Games
- asymetria graczy:
  - obrońca (lider, leader)
  - atakujący (naśladowca, follower)

- Atakujący zna strategię Obrońcy w momencie wyboru swojej strategii
- Atakujący rozstrzyga remisy na korzyść Obrońcy

Gry obronne Stackelberga

- Dwupoziomowy problem optymalizacyjny:

$$\arg\max_{\sigma_d \in \Sigma_d} U_d(\sigma_d, R_a(\sigma_d))$$

$$R_a = \arg\max_{\sigma_a \in \Sigma_a} U_a(\sigma_d, \sigma_a)$$

Gry obronne Stackelberga - przykład

		Atakujący	
		Cel 1	**Cel 2**
Obrońca	**Cel 1**	4, -3	-1, 1
	Cel 2	-5, 5	2, -1

Strategie proste:

obrońca cel 1, atakujący cel 2.

Wypłata atakującego: 1

Wypłata obrońcy: -1

Gry obronne Stackelberga - przykład

		Atakujący	
		Cel 1	**Cel 2**
Obrońca	**Cel 1**	4, -3	-1, 1
	Cel 2	-5, 5	2, -1

Strategie mieszane:

Obrońca: cel 1 - 50%, cel 2 – 50%

Atakujący atakuje cel 1: 0,5*(-3)+0,5*5 = 1

Atakujący atakuje cel 2: 0,5*1+0,5*(-1) = 0

Zatem atakujący wybiera cel 1 i wtedy oczekiwana wypłata obrońcy to: 0,5*4+0,5*(-5) = -0.5

Gry obronne Stackelberga - przykład

		Atakujący	
		Cel 1	**Cel 2**
Obrońca	**Cel 1**	4, -3	-1, 1
	Cel 2	-5, 5	2, -1

Strategie mieszane:

Obrońca: cel 1 - 60%, cel 2 – 40%

Atakujący atakuje cel 1: 0,6*(-3)+0,4*5 = 0,2

Atakujący atakuje cel 2: 0,6*1+0,4*(-1) = 0,2

Zatem atakujący wybiera cel z lepszą wypłatą obrońcy:

Wypłata obrońcy **atak na cel 1**: 0,6*4+0,4*(-5) = 0,4

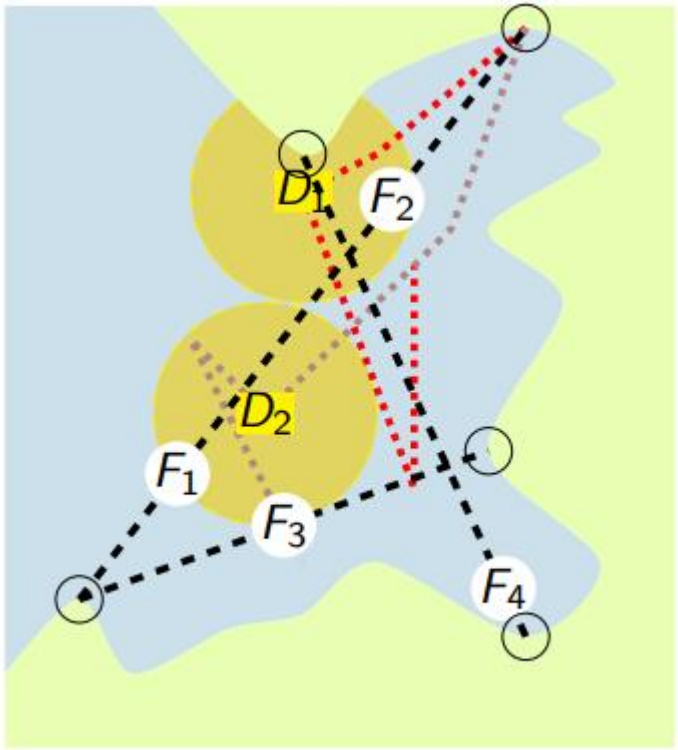
Wypłata obrońcy atak na cel 2: 0,6*(-1)+0,4*2 = 0,2

Algorytm ewolucyjny

- problem znalezienia równowagi Stackelberga jako problem optymalizacyjny
- przestrzeń przeszukiwań: strategie mieszane obrońcy
- funkcja celu: oczekiwana wypłata obrońcy
- ewaluacja rozwiązań: wypłata obrońcy przy optymalnej odpowiedzi atakującego

Ogólny schemat algorytmu ewolucyjnego

Gra z ruchomymi celami



- m dyskretnych kroków czasowych
- F_j – promy poruszające się zgodnie z zadanym planem
- $D_1, D_2, \ldots, D_{nd}$ – jednostki obrońcy
- atakujący wybiera jeden z promów w dowolnym kroku czasowym lub rezygnuje z ataku
- prom jest chroniony, jeśli w odległości co najwyżej r znajduje się jednostka obrońcy
- atak trwa k jednostek czasu
- atak jest nieudany jeśli w dowolnym momencie ataku prom jest chroniony

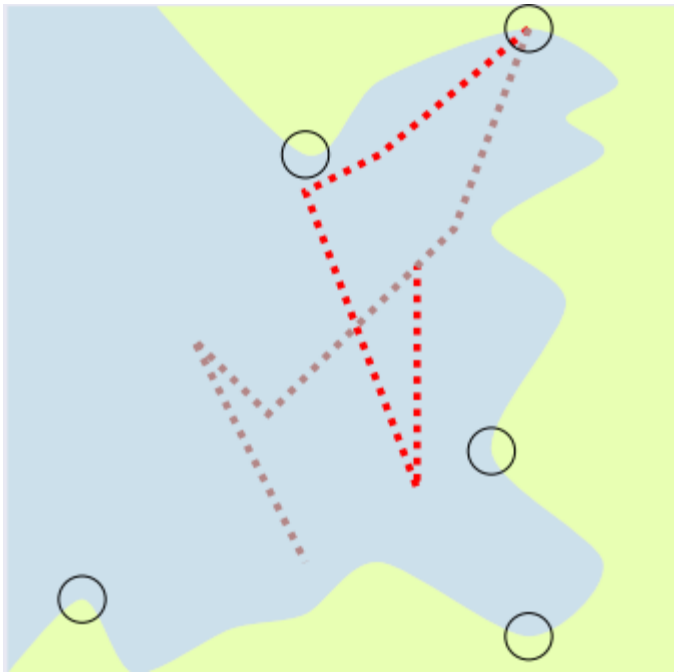
Kodowanie rozwiązań

- strategia prosta:

$$SS = [(D_1(t_1), D_1(t_2), \ldots, D_1(t_m)), \ldots, (D_2(t_1), \ldots, D_2(t_m))]$$

- chromosom – zbiór strategii prostych z prawdopodobieństwami

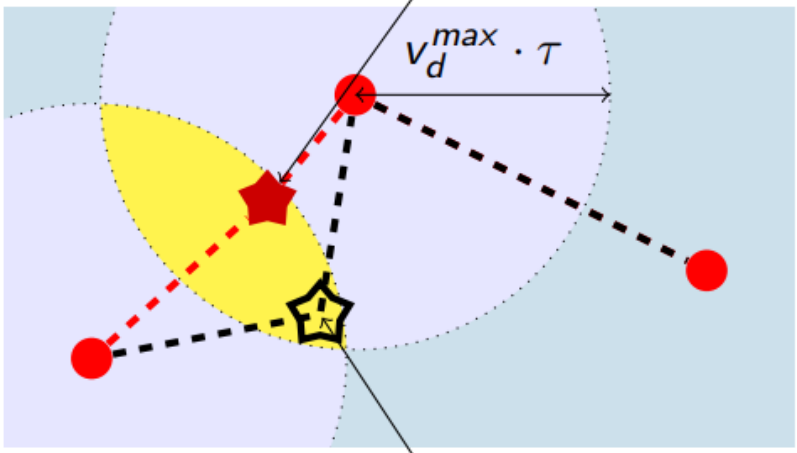
$$CH_q = \{(SS_1^q, p_1^q), \ldots, (SS_{l_q}^q, p_{l_q}^q)\}, \quad \sum_{i=1}^{l_q} p_i^q = 1$$



Mutacje

- modyfikacja pozycji obrońcy w losowym interwale czasowym w losowej strategii prostej
- założenie: brak konieczności modyfikacji pozostałych pozycji

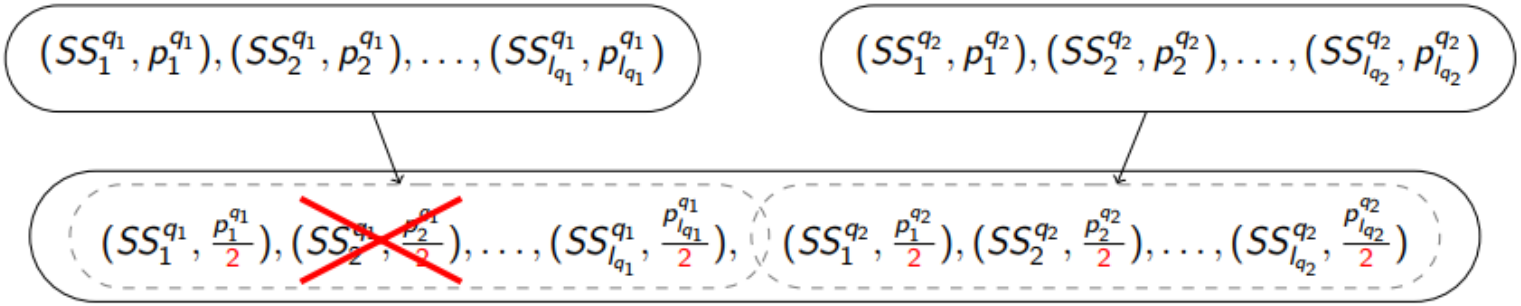
$$CH = [\ldots, (p_i, (D_1(t_1), \boxed{D_1(t_2)}, D_1(t_3), D_1(t_4)), \ldots), \ldots]$$



$$CH' = [\ldots, (p_i, (D_1(t_1), D_1'(t_2), D_1(t_3), D_1(t_4)), \ldots), \ldots]$$

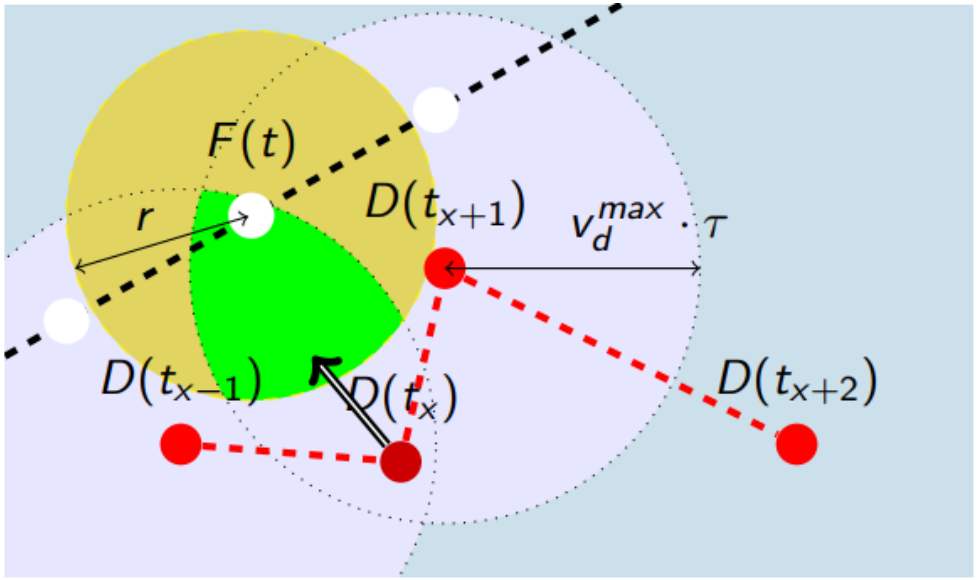
Krzyżowanie

- połączenie dwóch strategii mieszanych w jedną – prawdopodobieństwa dzielone na pół
- usunięcie losowych strategii prostych z prawdopodobieństwem odwrotnym do prawdopodobieństwa wyboru danej strategii



Lokalna optymalizacja

- przesunięcie jednostek obrońcy w chwili t tak, aby:
  - nie zmienić pozycji w chwilach t-1 oraz t+1
  - ochraniać jak największą liczbę celów

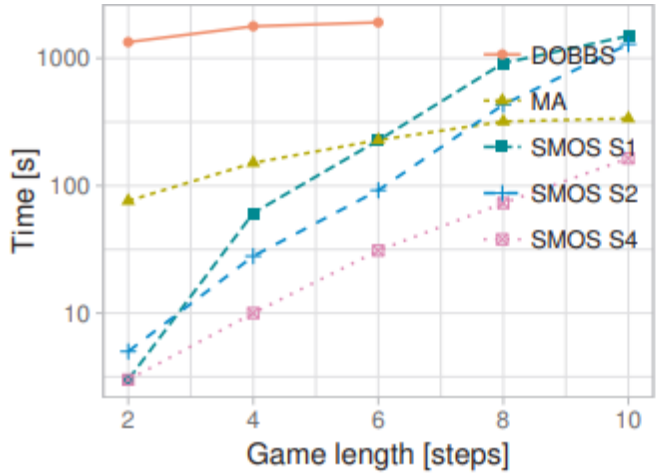


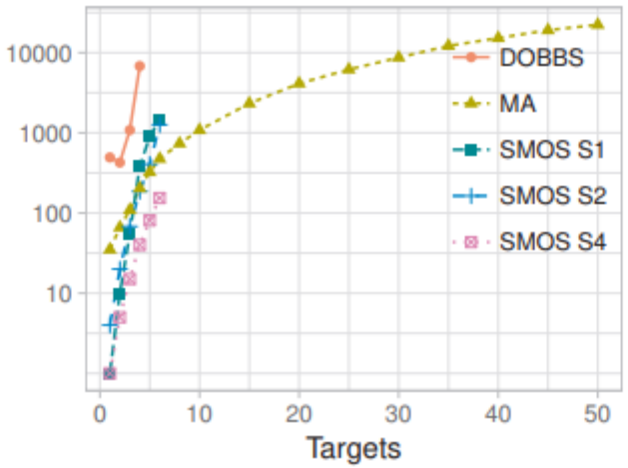
Parametry

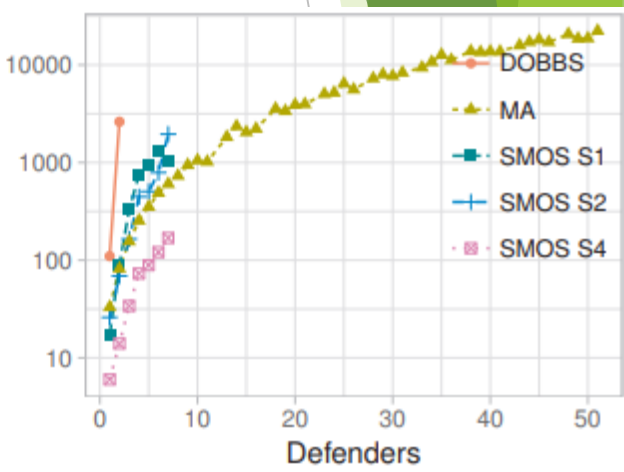
parametr	wartość
wielkość populacji	2000
liczba iteracji	2000
pstwo mutacji	0.2
powtórzeń mutacji	10
pstwo krzyżowania	0.9
rodzaj selekcji	turniejowa binarna
presja selekcji	0.9
elita	2

Wyniki

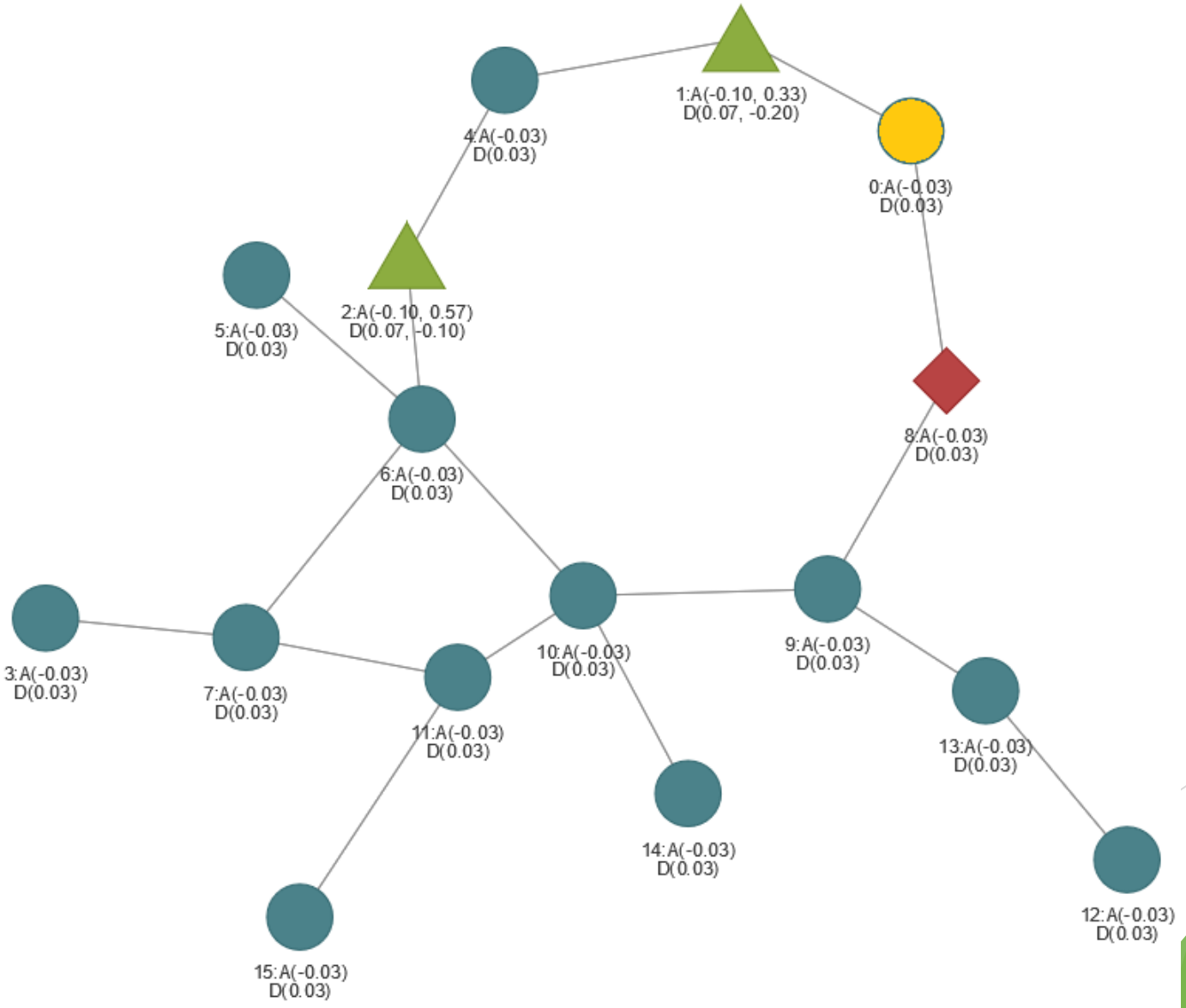
- lepsza skalowalność
- przeszukiwanie przestrzeni ciągłej
- optymalny wynik znaleziony dla 599 z 605 gier testowych







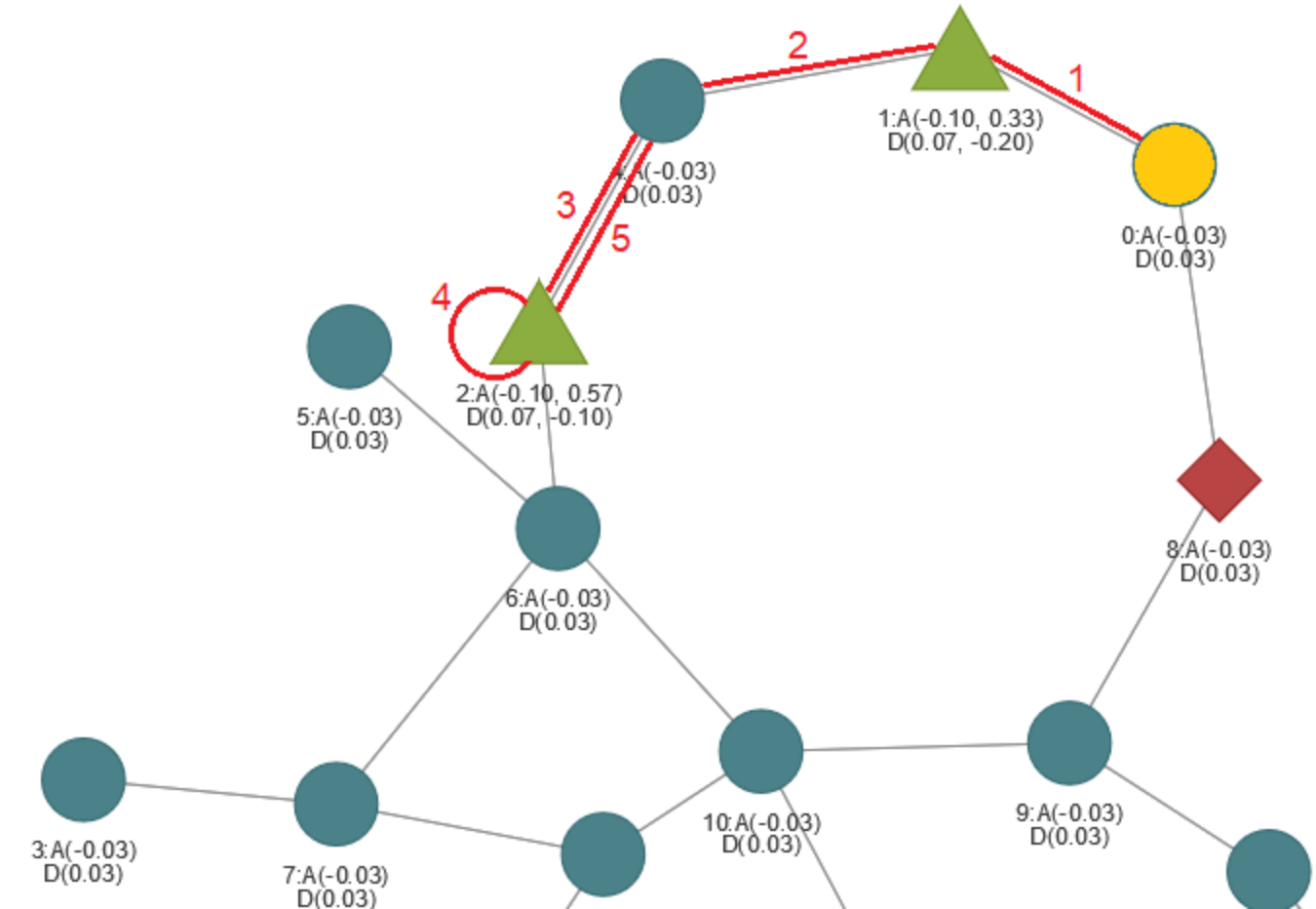
Gry na grafie



Kodowanie rozwiązań

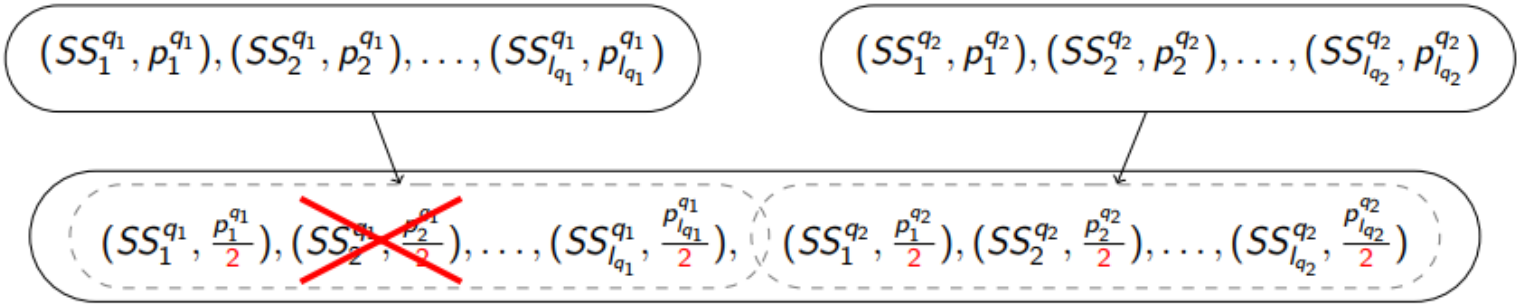
- zbiór strategii prostych z prawdopodobieństwami

$$CH_q = \{(SS_1^q, p_1^q), \ldots, (SS_{l_q}^q, p_{l_q}^q)\}, \quad \sum_{i=1}^{l_q} p_i^q = 1$$



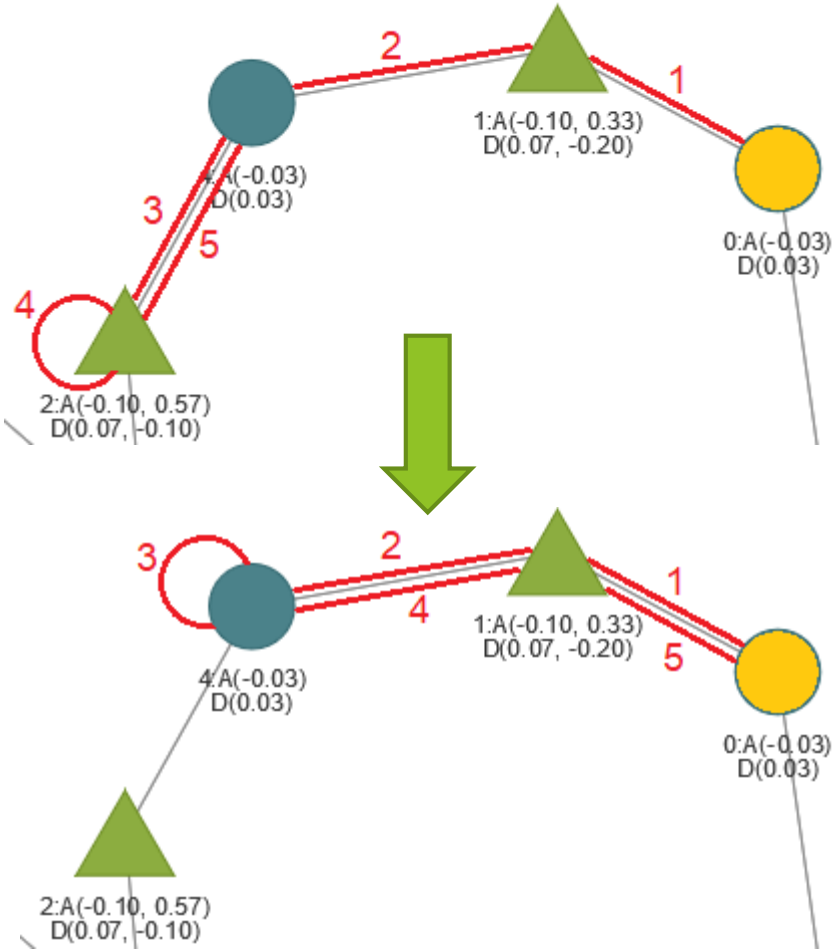
Krzyżowanie

- połączenie dwóch strategii mieszanych w jedną – prawdopodobieństwa dzielone na pół
- usunięcie losowych strategii prostych z prawdopodobieństwem odwrotnym do prawdopodobieństwa wyboru danej strategii



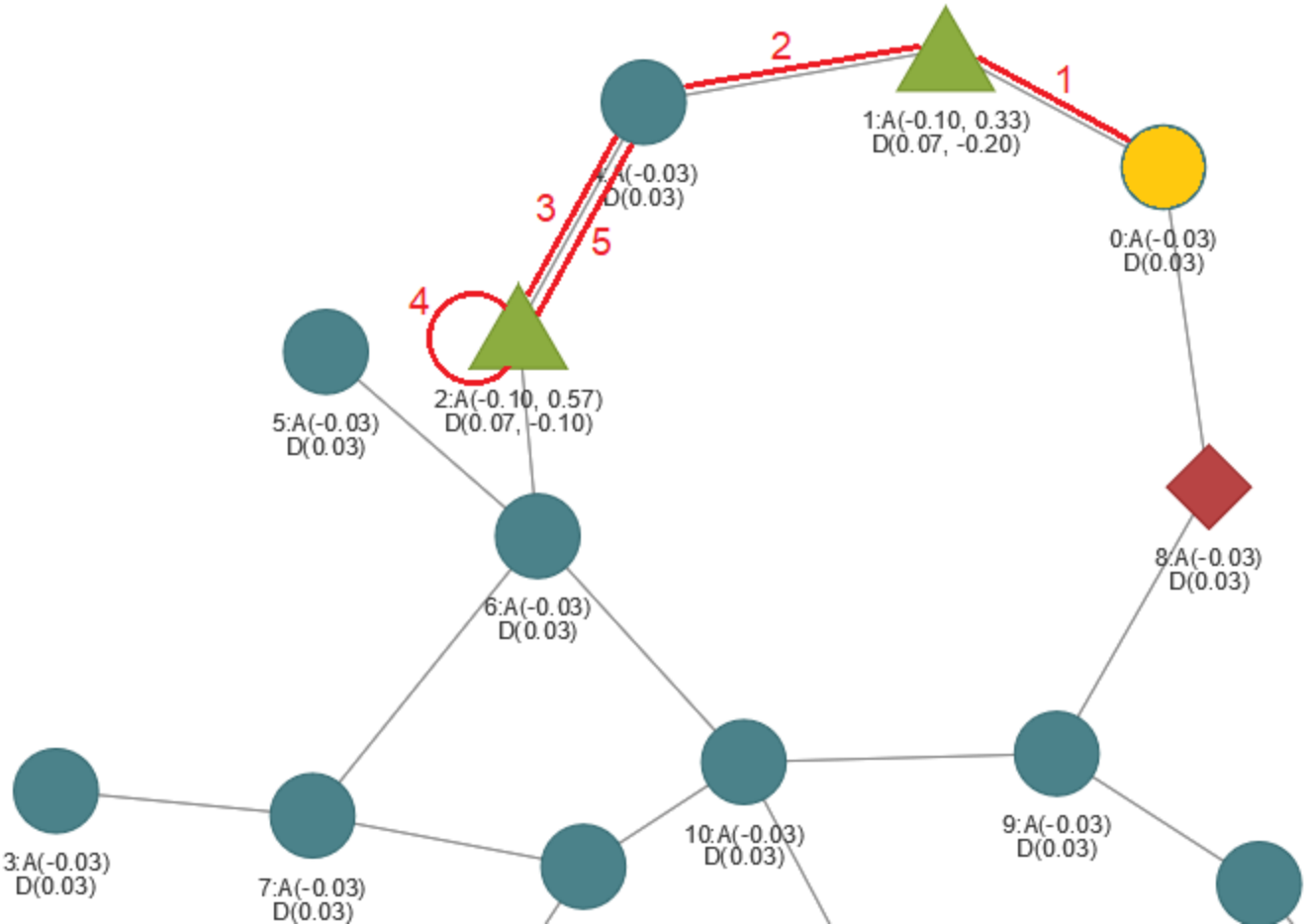
Mutacje

- modyfikacja pozycji obrońcy od losowego interwału czasowego w losowej strategii prostej



Lokalna optymalizacja

- BRAK

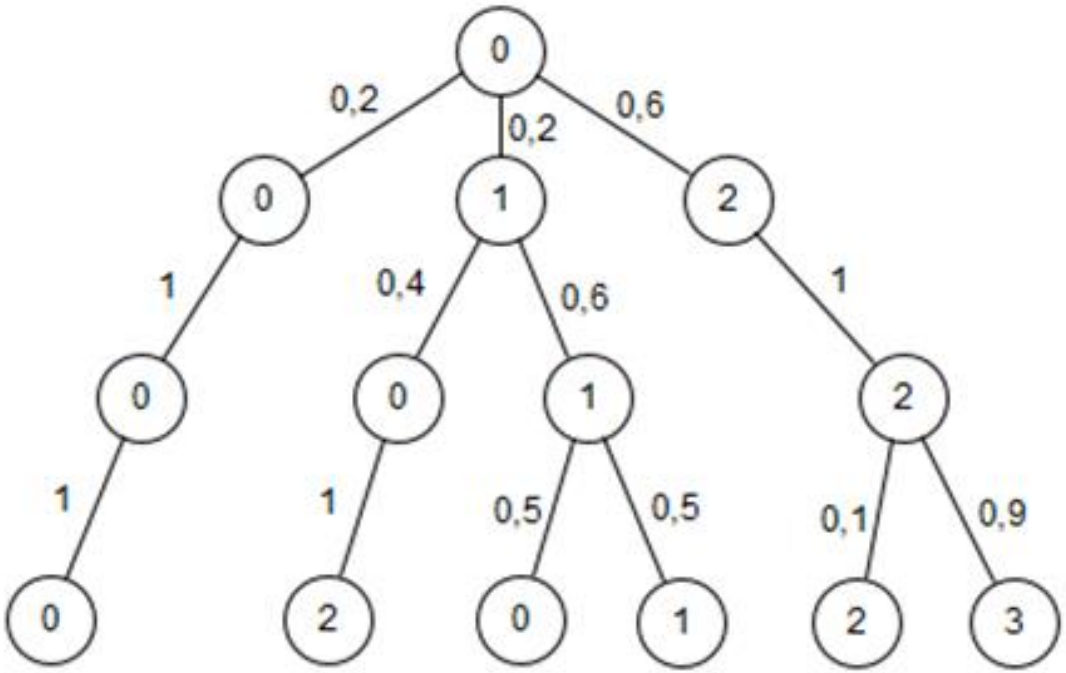


Wyniki

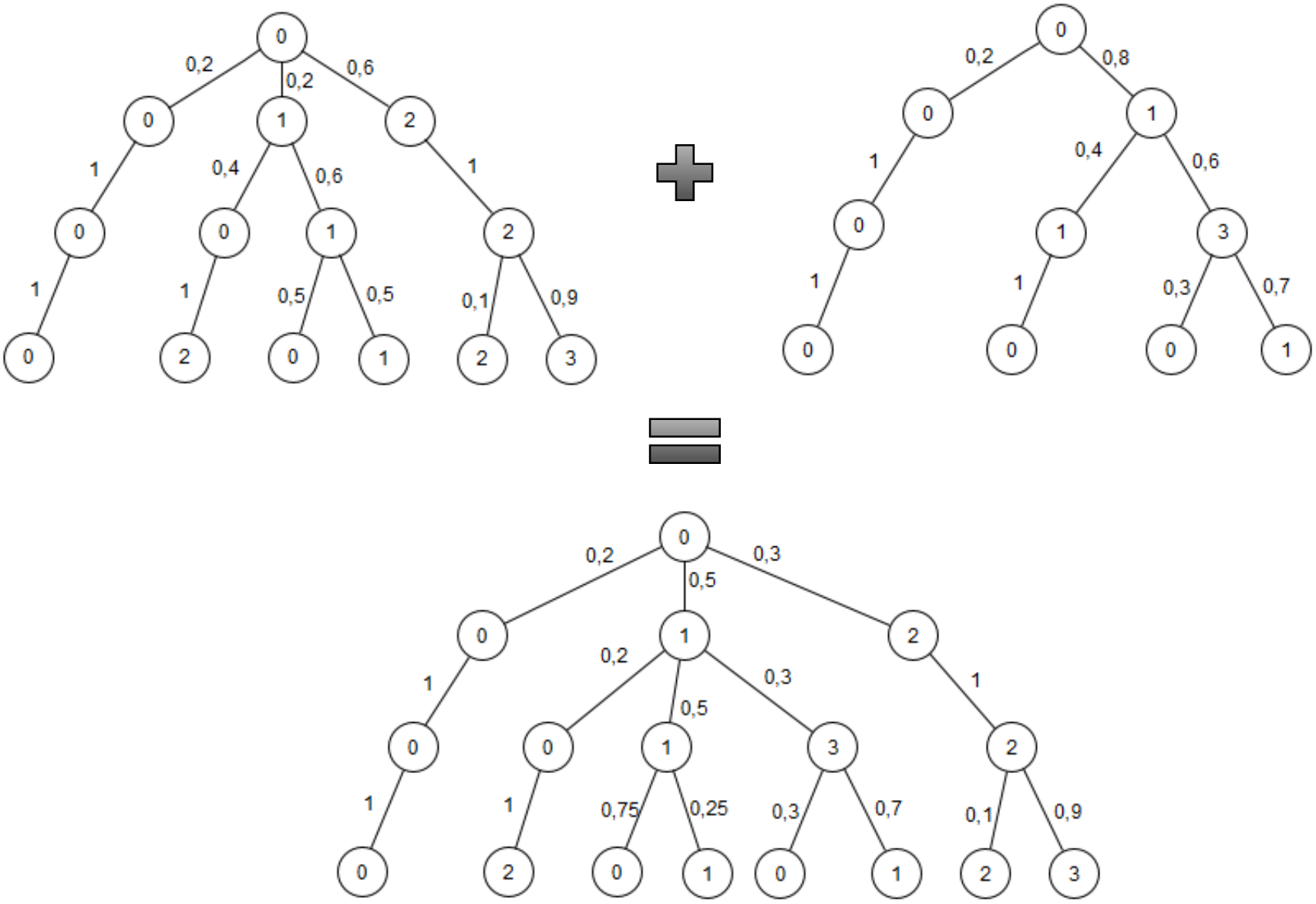
- 50 gier 5- i 6-krokowych
- w 44 przypadkach znalezione optymalne rozwiązanie
- znacznie szybszy czas działania:
  - rozwiązanie dokładne: ~2h
  - zaproponowany algorytm ewolucyjny: ~45s

Alternatywna reprezentacja

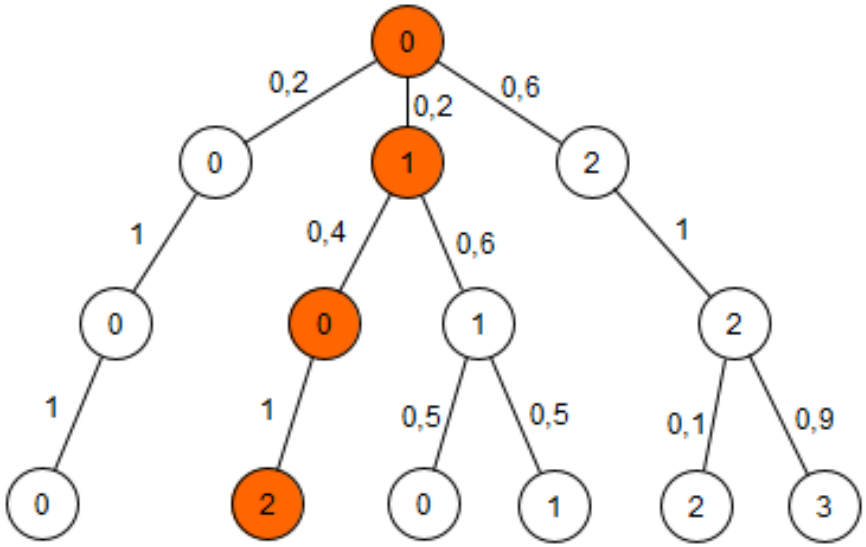
- Drzewo strategii

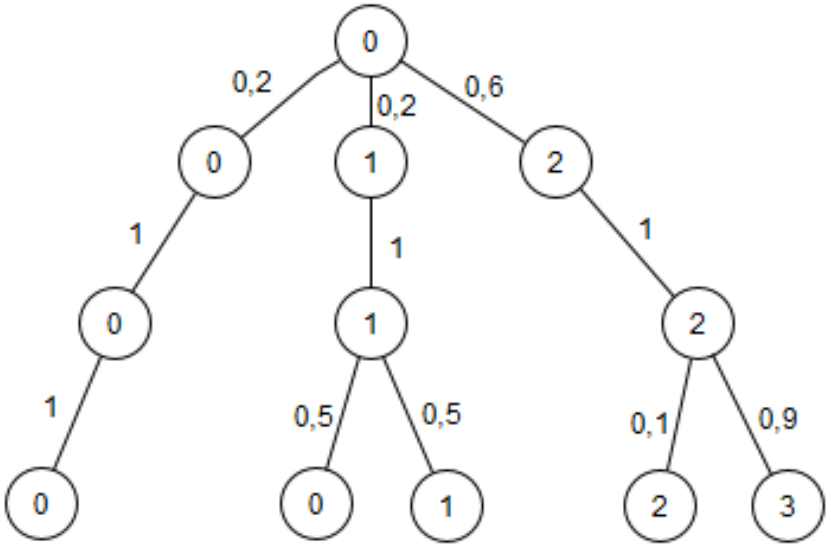


Krzyżowanie

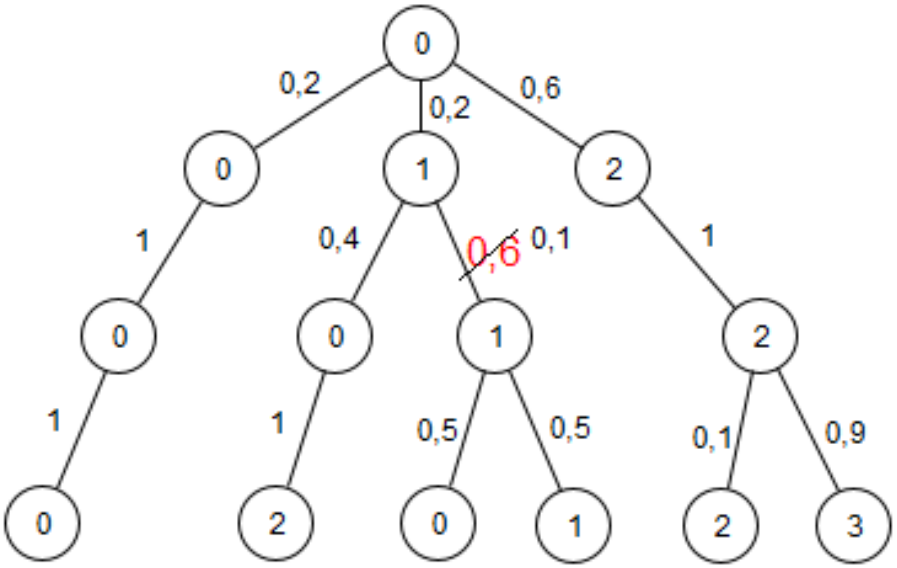


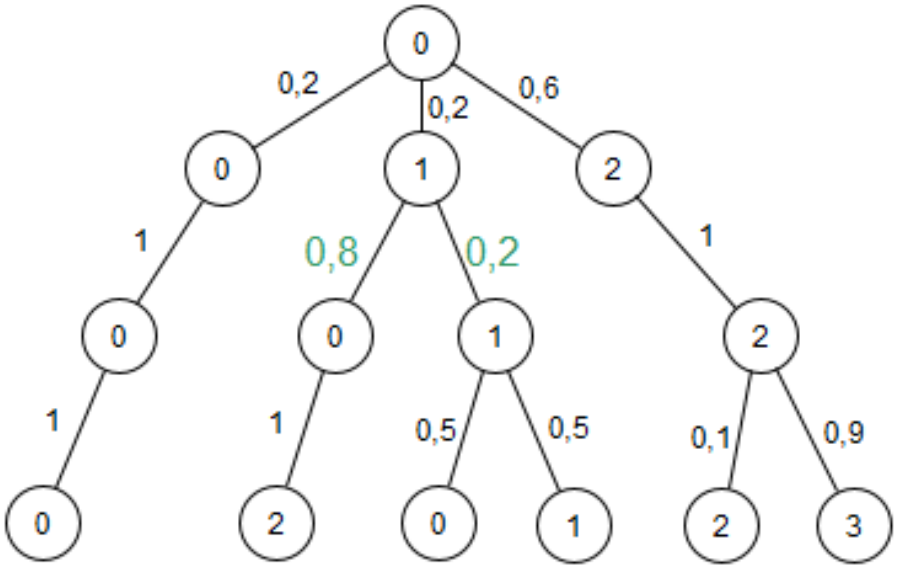
Mutacja 1 – usuwanie strategii



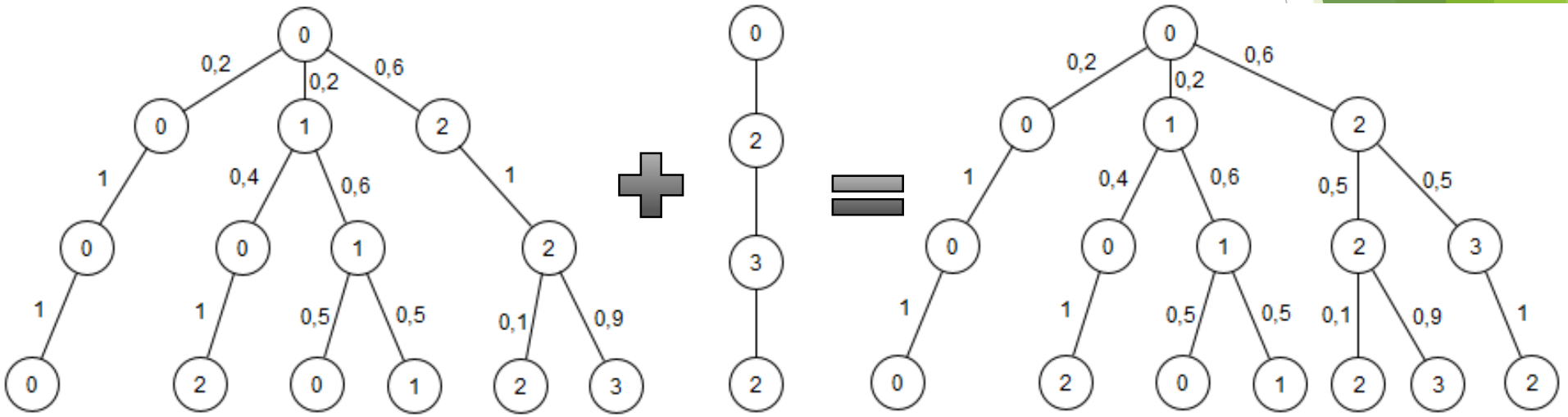


Mutacja 2 – zmiana prawdopodobieństwa





Mutacja 3 – dodanie strategii



Wyniki

- nieznacznie gorsze wyniki (optymalne rozwiązanie znalezione dla 42 z 50 gier)
- dłuższy czas działania: ~60s

Dalszy kierunek badań – algorytm koewolucyjny

- rywalizujące populacje strategii obrońcy i atakującego
- ewaluacja strategii obrońcy nie wymaga sprawdzenia wszystkich strategii atakującego
- potencjalne zastosowanie do nowych problemów: z nieskończoną przestrzenią decyzyjną atakującego

Bibliografia

[1] J. Karwowski, J. Mańdziuk, A. Żychowski, F. Grajek, B. An, (in print, 2019), *A Memetic Approach for Sequential Security Games on a Plane with Moving Targets*, 33rd AAAI Conference on Artificial Intelligence (AAAI-2019).

[2] J. Karwowski, J. Mańdziuk, (2019), *A Monte Carlo Tree Search approach to finding efficient patrolling schemes on graphs*, European Journal of Operational Research, vol. 277, 255-268.

[3] X. Wang, et al. *Catching Captain Jack: Efficient time and space dependent patrols to combat oil-siphoning in international waters*. 32nd AAAI Conference on Artificial Intelligence. (AAAI-2018).

[4] J. Gan, et al. *Security games on a plane*. 31st AAAI Conference on Artificial Intelligence (AAAI-2017).

[5] P. Paruchuri, et al. *Playing games for security: An efficient exact algorithm for solving Bayesian Stackelberg games*. Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2. 2008.