Security Games: zastosowanie UCT na tle dotychczasowych podejść do problemu

Jan Karwowski

Zakład Sztucznej Inteligencji i Metod Obliczeniowych
Wydział Matematyki i Nauk Informacyjnych PW

14 X 2014

Security Games

Zastosowania

- Zabezpieczenie lotnisk przed zamachami
- Zabezpieczenie ruchomych celów
- Zabezpieczenie zasobów przed kradzieżą

Model miejsca rozgrywki

- Zbiór miejsc, bez zadanych relacji przestrzennych
- Graf
- Przestrzeń ciągła

Gracze

- Broniący
- Atakujący (jeden lub wielu)

Pursuer-Evader Game

Gracze

- Ścigany
- Ścigający — cel: znalezienie się w tym samym miejscu i czasie co Ścigany
- Grają przeciw sobie

- Czas gry może nie być ściśle określony
- Może się rozgrywać na grafie (znalezienie się w tym samym wierzchołku) lub w przestrzeni ciągłej (znalezienie się bliżej niż ϵ w pewnej metryce)

Pursuer-Evader Game

Przykład (Example 3)

Kobieta pływa w jeziorze w kształcie koła z szybkością v_l. Mężczyzna, który chciałby przeszkodzić kobiecie, niestety nie umie pływać. Może on biegać dookoła jeziora z szybkością v_m. Kobieta biega szybciej ob mężczyzny i jeśli wyjdzie z wody, ten nie ma szans jej dogonić. Kobieta chce w końcu opuścić jezioro i jednocześnie nie chce spotkać się z mężczyzną. Każde z nich wykonuje zmiany kierunku swojego ruchu, aby osiągnąć swój cel.

T. Basar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. Academic Press Inc., 1982

Gra macierzowa

- Dwóch graczy
- Dwie macierze A – wypłaty pierwszego gracza, B — drugiego
- Wiersze macierzy etykietowane są ruchem (sekwencją ruchów) pierwszego gracza
- Kolumny etykietowane ruchem (sekwencją) drugiego gracza
- Jeśli gra o sumie zerowej, to $A = -B$

Gra Stackelberga

Wypłaty

Gra nie ma sumy zerowej!

Gracze

- Leader
- Follower

Równowaga Stackelberga

- Follower zna taktykę leadera
- Follower gra w pełni racjonalnie, maksymalizując swoją wygraną
- Potrzeba rozwiązania Dyskretno-ciągłego problemu optymalizacji liniowej (MILP)

Security Games

- Niewspółpracujący ze sobą gracze
- Bardzo często gra typu Pursuer-Evader
- Rozwiązania często bazują na modelu Stackelberga
- **Celem jest znalezienie możliwie najlepszej strategii dla *broniącego***
- Niepełna informacja

Wdrożone rozwiązania

ARMOR Assistant for Randomized Monitoring over Routes – planowanie rejonów patroli na lotnisku LAX (2007) [2]

IRIS Inteligent Randomization in Scheduling – rozmieszczanie tajnych agentów w rejsowych samolotach (2009) [2]

PROTECT Port Resilience Operational/Tactical Enforcement to Combat Terrorism – planowanie patroli US Coast Guard (2011) [4]

Model ARMOR (LAX)

- Planowanie punktów kontrolnych na drogach dojazdowych do lotniska
- Planowanie patroli z psami na terenie terminali

Gra Stackelberga

- Abstrakcyjna przestrzeń punktów, nie uwzględnia rozmieszczenia przestrzennego lotniska
- Arbitralnie dobrane wypłaty dla atakujących i broniących w zależności od rozkładu lotów w danej godzinie
- Modyfikacje problemu optymalizacyjnego

Ograniczona racjonalność ruchów przeciwnika

Problem

W rzeczywistych warunkach bojowych atakujący są pod dużą presją psychiczną i ich zachowanie może nie być do końca racjonalne. Przeczy to założeniu Równowagi Stackelberga, że ruchy Followera są optymalne.

Ponadto w rzeczywistości nie jest prawdziwe założenie, że Follower zna pełną strategię Broniącego.

- QR – Quantal Response [3] — wypłaty widziane przez przeciwnika są zaburzone o niewielki losowy składnik
- Systematyczne zaburzenia (wartości oczekiwanych) wypłat związane z ograniczeniami i zaburzeniami percepcji
- Systematyczne zaburzenia optymalizowanych funkcji

Definicja gry I

- $G = (V, E)$ – graf skierowany, V – zbiór, $E \subseteq V \times V$
- $T \subset V$ – cele
- $S \subset V$, $S \cap T = \emptyset$ – wierzchołki startowe atakujących
- $\tau \in \{1, 2, \ldots\}$ – bieżący krok czasowy gry
- $a \in A$ – atakujący, zbiór atakujących
- $u \in U$ – jednostki, którymi dysponuje gracz broniący
- $\mathbb{S}$ – zbiór wszystkich możliwych stanów gry
- $R_A : A \times T \times \mathbb{S} \to \mathbb{Z}^+$ – wypłata za skuteczny atak na cel
- $P_D : A \times T \times \mathbb{S} \to \mathbb{Z}^-$ – kara dla broniącego, za nieskuteczną obronę celu
- $R_D : A \times V \times \mathbb{S} \to \mathbb{Z}^+$ – wypłata broniącego za złapanie atakującego
- $P_A : A \times V \times \mathbb{S} \to \mathbb{Z}^-$ – kara atakującego za bycie złapanym

Definicja gry II

- $p : V \times \mathbb{N} \times \mathbb{S} \rightarrow [0, 1]$ – prawdopodobieństwo, że atakujący zostanie złapany, jeśli na tym samym wierzchołku jest zadana liczba obrońców

Czas

- Dyskretne jednostki czasu
- W każdym kroku czasowym jednostki atakujące i broniące mogą przejść jedną krawędź wychodzącą z wierzchołka i znaleźć się w jej drugim końcu
- Po dokonaniu jednoczesnych przesunięć atakujących i broniących następuje sprawdzenie czy atakujący i broniący znajdują się na tym samym wierzchołku i sprawdzenie, czy atakujący został złapany

Definicja gry III

Uwagi

- Wypłaty mogą się różnić w zależności od stanu gry
- Atakujący może mieć różną liczbę jednostek (ograniczoną lub nie)
- Do obrony celów potrzeba więcej jednostek niż ma do dyspozycji broniący

Stan gry $\overline{S} \in \mathbb{S}$

- $L_D : U \rightarrow V$ – położenia jednostek obrony
- $D_D : U \rightarrow V \cup \overline{0}$ – cele przemieszczeń jednostek obrony
- $L_A : A \rightarrow \{(i, v) | i \in \mathbb{N} \wedge v \in V\}$ – położenia jednostek atakujących (ich liczba na planszy może się zmieniać)
- $D_A : A \rightarrow \{(i, v) | i \in \mathbb{N} \wedge v \in V\}$ – cele przemieszczeń atakujących
- Opcjonalnie: historia incydentów
- Informacje charakterystyczne dla danej gry, np. rozkład lotów

Nie jest to gra o pełnej informacji. Każdy z graczy widzi jakąś funkcję stanu gry, która nie zawiera wszystkich elementów stanu, w szczególności bieżących pozycji przeciwnika.

Zastosowanie UCT

- Nieskończona gra
  - Koniec, gdy $f(div, \tau)$ będzie odpowiednio duże. div – liczba odwiedzonych stanów w tej rozgrywce (różnorodność rozgrywki). Aktualnie:
$$f(div, t) = 3 * div + \log t$$
- Wypłaty w różnych momentach
  - Aktualizowanie wypłat w węźle według malejącej funkcji odległości czasowej od wypłaty
- Węzły etykietowane stanem gry widzianym przez broniącego
- Ruchami w grze są polecenia zmiany ustawienia jednostek

Ruchy obrońcy I

Wariant 1

Broniący może wykonywać następujące ruchy (dodatkowo jest ograniczenie na wykonalność, ruch przesunięcia można wykorzystać tylko gdy jest jednostka obrony, którą można w ten sposób przesunąć):

- Zleć przesunięcie jednostki z wierzchołka v_i do v_j, $j \neq i$.
- Zakończ przesunięcia i czekaj na kolejny krok czasowy.

W przypadku, gdy w punkcie jest więcej niż jeden patrol, przesuwany jest ten o niższym numerze porządkowym.

Ruchy obrońcy II

Wariant 2*

Wariant podobny do Wariantu 1, z tym że przesunięcie definiowane jest dla jednostki o konkretnym numerze. Dostępne ruchy:

- Zleć przesunięcie jednostki u_i do wierzhołka v_j.
- Zakończ przesunięcia i czekaj na kolejny krok czasowy.

Ruch przesunięcia jest dopuszczalny tylko gdy jednostka nie dostała rozkazu przesunięcia do tego punktu w tym kroku czasowym.

Ruchy obrońcy III

Wariant 3

Jest dostępna tylko jedna rodzina ruchów, wykonanie ruchu powoduje przesunięcie kroku czasowego.
Ruch polega na zleceniu nowego oczekiwanego ustawienia jednostek obrony..
Za decyzję które jednostki będą przesuwane będzie odpowiedzialny prosty algorytm (przy założeniu, że jednostek obrony jest kilka, można założyć że jesteśmy w stanie dla tej wielkości rozwiązać problemy NP-trudne). Do rozważenia:

- Minimalizacja czasu kiedy ostatnia jednostka dotrze na miejsce
- Minimalizacja liczby przesuwanych jednostek
- Minimalizacja sumarycznego czasu przemieszczania się wszystkich

Potrzeba istnienia przeciwnika

UCT w fazie nauki musi rozegrać wiele gier z przeciwnikiem.

Przeciwnik

- Zbieranie częstości pojawiania się obrońców w poszczególnych wierzchołkach
- Ograniczanie racjonalności
  - Systematyczne
  - Losowe

Można użyć wielu przeciwników w jednej rozgrywce.

Problemy

- Wiele ruchów w kroku czasowym: czy dopuścić zmianę celu jednostki (2)? Karanie za zbyt wiele ruchów w kroku?
- Jak i z czym porównywać wyniki?

Literatura I

[1] T. Basar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. Academic Press Inc., 1982.

[2] Manish Jain et al. "Software assistants for randomized patrol planning for the lax airport police and the federal air marshal service". In: *Interfaces* 40.4 (2010), pp. 267–290.

[3] Richard D. McKelvey and Thomas R. Palfrey. "Quantal Response Equilibria for Normal Form Games". In: *Games and Economic Behavior* 10.1 (1995), pp. 6 –38. ISSN: 0899-8256. DOI: http://dx.doi.org/10.1006/game.1995.1023. URL: http://www.sciencedirect.com/science/article/pii/S0899825685710238.

Literatura II

[4] Eric Shieh et al. “PROTECT: A deployed game theoretic system to protect the ports of the United States”. In: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems. 2012, pp. 13–20.