Posudek diplomové práce

Autor: Richard Klíma

Název: Combining Online Learning and Equilibrium computation in Security Games

Posudek vypracoval vedoucí práce: Ing. Ondřej Vaněk, Ph.D., Katedra počítačů, FEL, ČVUT

Cílem této diplomové práce bylo zkombinovat algoritmy pro výpočet equilibrií v bezpečnostních hrách (security games) a algoritmy pro on-line učení v tzv. adversarial setting. Tzn. nastudovat literaturu relevantní těmto směrům, empiricky evaluovat výkon těchto algoritmů proti různým modelům oponenta, navrhnout kombinovanou metodu, která začíná pracovat s nepřesným herně-teoretickým modelem a postupně se zlepšuje na základě opakované interakce s oponentem a ve finále prostudovat aplikovatelnost těchto algoritmů v problémech z reálného světa.

Autor zadání diplomové práce splnil. V první části práce pečlivě prezentuje jednotlivé metody a vysvětluje algoritmy relevantní jeho práci. Soustředí se na Stackelbergovy bezpečnostní hry a na výpočet equilibrií v těchto hrách. Dále studuje učící algoritmy pro multi-armed bandit problém, primárně UCB A EXP3 algoritmy a jejich variace, které přehledně popisuje.
V druhé části práce student definuje typy chování hráčů, tj. obránce a útočníka a definuje variace Stackelbergových strategií při neúplné informaci o protistraně. Zde stojí za zmínku design tzv. motivovaného Stackelbergova equilibria (SSE) vylepšující hru proti útočníkovi s nepřesnou informací. Dále student prezentuje jím vytvořené čtyři kombinace algoritmů, které kombinují herně-teoretický přístup a přístup založený na on-line učení, specificky strategii získanou z Nashova equilibria s EXP3 algoritmem. Student také prezentuje kombinatorické verze algoritmů, které uvažují více jednotek obránce.
Ve třetí fázi student provádí rozsáhlou sadu experimentů, které hledají optimální nastavení parametrů pro jednotlivé algoritmy, demonstrují vlastnosti studovaných algoritmů na sadě protivníkovo modelů, ukazují slabinu tzv. “fictious play” modelu útočníka, který je exploitovatelný UCB algoritmem, demonstrují efektivitu motivovaného Stackelbergova equilibria a dále obšírně zkoumají vlastnosti kombinovaných algoritmů.

Student zpracoval dané téma velmi pečlivě, i když v relativně specifickém rámci, tj. kombinuje EXP3 algoritmus s Nashovo equilibriem. Zde bych očekával větší variaci v designu a kombinaci EXP3/UCB s NE i SSE. Zadání také mluví o studiu aplikovatelnosti přístupů vzhledem k praktickým aplikacím - v textu je zmíněný problém patrolování hranic, nicméně student nijak nekomentuje vlastnosti navržených algoritmů v tomto specifickém problému. Hledání hodnot parametrů algoritmů je provedeno velmi pečlivě a nemám k tomu námitek, bohužel v dalších kapitolách se používají jiné hodnoty (viz sekce 5.1.2). Vysvětlení exploitovatelnosti “fictious play” modelu oponenta je také rozebráno velmi pečlivě, nicméně i přesto, že je model exploitovatelný učícími strategiemi (zde demonstrováno na UCB), se tento

model používá jako benchmark ve velké množině evaluačních scénářů. Model útočníka se změnami je relativně jednoduchý a demonstrace špatného výkonu SSE strategie proti učícím algoritmům velmi umělý. Zde bych ocenil model útočníka s graduální změnou a porovnání učících algoritmů s přepočítáváním SSE s různou periodou.
Po formální stránce je práce zpracována nadprůměrně, text je pečlivě napsaný, neobsahuje gramatické chyby, práce má správnou strukturu a student pracuje správně s literaturou. Z tohoto pohledu nemám k práci námitek.

Při obhajobě doporučuji autorovi položit následující doplňující dotazy:

1. dotaz 1: Demonstrujete exploitaci “fictious play” modelu strategií založenou na UCB. Dokáže také EXP3 (či modifikace tohoto algoritmu) “fictious play” model exploitovat?
2. dotaz 2: V zadání se mluví o postupném **zpřesňování strategie** vypočtené na základě nepřesného modelu na základě opakované interakce s oponentem. Jak byste přistoupil k **zpřesňování herně-teoretického modelu** na základě interakce s oponentem?

Předloženou diplomovou práci hodnotím známkou **B-velmi dobře.**

V Praze dne 25. 5. 2015 Ing. Ondřej Vaněk Ph.D.