

Warszawa, 14 marca 2019

dr hab. inż. Szymon Jaroszewicz  
Instytut Podstaw Informatyki PAN  
ul. Jana Kazimierza 5  
Warszawa

## Recenzja rozprawy doktorskiej mgr. Dominika Deji

Promotor: dr hab. Adam Wierzbicki, prof. PJATK  
Promotor pomocniczy: dr inż. Radosław Nielek

### Uwagi ogólne

Rozprawa doktorska Pana mgr. Dominika Deji ma formę zwartej publikacji. Dotyczy ona zastosowań *data science* w e-sporcie, czyli profesjonalnym współzawodnictwie w grach komputerowych.

Pan Deja przedstawił trzy cele badawcze, dotyczące rozważanego tematu, a mianowicie: (1) opracowanie algorytmu łączenia graczy w drużyny, (2) analizę rozgrywek przy pomocy wzorców czasowo przestrzennych oraz (3) aproksymację funkcji wartości gier w warunkach zmieniającej się w czasie struktury gry. Struktura pracy podporządkowana jest owym celom, które autor po kolei analizuje.

Przejdę teraz do oceny wkładu pracy w dziedzinę informatyki a następnie przedstawię merytoryczną oceną uzyskanych wyników.

### Ocena wkładu pracy w dziedzinę informatyki

Praca dotyczy analizy gier komputerowych metodami uczenia maszynowego i optymalizacji. Metody te są bardzo blisko związane z informatyką, komponent informatyczny jest zatem w pracy bardzo istotny. Dodatkowo, zastosowanie tych metod nie jest trywialne i wymaga ich rozwinięcia co stanowi

wkład w rozwój tych metod. Przykładem niech będzie metoda oceny dokładności modeli klasyfikacyjnych zaproponowana przez autora w rozdziale 6, która jest twórczą modyfikacją istniejących metod dostosowującą je do przypadku gier komputerowych ze zmieniającymi się zasadami i jednocześnie do innych problemów o podobnej strukturze. Nie mam zatem wątpliwości, że pracę należy uznać za pracę z dziedziny informatyki.

## Ocena merytoryczna uzyskanych wyników

Przejdę teraz do omówienia kolejnych rozdziałów pracy i oceny ich istotności naukowej oraz wkładu w dziedzinę informatyki społecznej.

Pierwszym rozważanym przez autora problemem jest przydział graczy do pięcioosobowych drużyn w grze League of Legends.

Ważnym elementem pracy doktoranta jest uwzględnienie problemu tzw. *toxicity* czyli nieprzyjaznych zachowań wśród członków drużyny. Problem ten nie był uwzględniany we wcześniejszych podejściach. Osiągnięto to, przez uwzględnienie różnych ról, które mogą być przypisane poszczególnym graczom.

Autor rozpoczął od pobrania danych rozgrywek. Problem ten nie zawsze był prosty, np. w przypadku gry League of Legends (LOL) trudnościami były ograniczenia w ilości informacji pobieranej przy użyciu oficjalnego API. Autor opracował specjalny algorytm (Algorytm 1) pozwalający na pobranie dużego a zarazem spójnego zbioru danych, co było zadaniem pracochłonnym i z pewnością nietrywialnym.

W części 4.2.1 autor analizuje wpływ korzystania z tzw. ‘najefektywniejszej strategii’ (oznaczanej skrótem META) na wynik gry i poziom frustracji graczy. Niestety autor nie wyjaśnia w jaki sposób ustalano czy dana drużyna stosuje META (którą autor definiuje jako korzystanie z pięciu różnych ról), czy informacja ta jest to dostępna w API?

W części 4.2.1.3 autor przedstawił interesującą analizę wpływu roli przypisanej graczowi na wynik rozgrywki. Stosując model logistyczny doktorant potwierdził, duży wpływ przypisania graczy do ich preferowanych ról i oszacował go ilościowo. Stanowi to zarazem wstęp i solidną motywację dla systemu łączenia graczy w drużyny zaproponowanego przez autora.

Formalna definicja zagadnienia opiera się na modyfikacji podejścia Almana i McKaya z pracy [5] poprzez wprowadzenie osobnych wartości umiejętności (*skill*) dla każdej roli gracza oraz jednoczesny przydział graczy do wielu meczy. Modyfikacja jest dość prosta, jednak spełnia zadanie postawione przez autora: uwzględnia preferencje graczy dotyczące ich ról w strategii (patrz jednak niżej). Dodatkowo, przydział do wielu rozgrywek na raz komplikuje nieco problem optymalizacyjny.

Problem przydziału graczy do drużyn według metody autora redukuje się do problemu programowania zero-jedynkowego, który jest problemem NP-zupełnym. Autor proponuje zatem interesujący algorytm przybliżony, który daje dobre wyniki eksperymentalne. Algorytm działa dwuetapowo, pierwszy etap maksymalizuje jednorodność drużyn, drugi dzieli grupy 10 graczy na dwie drużyny tak, by zrównoważyć ich względną siłę (*fairness*). Algorytm uznaję za elegancki przykład heurystycznej metody przybliżonej.

Niestety, przedstawione sformułowanie problemu optymalizacyjnego ma pewne wady. Po pierwsze, część ograniczeń jest zbędna. Np. ograniczenie 3 wynika z ograniczenia 2, a ograniczenie 4 z pozostałych ograniczeń. Nie jest to oczywiście problemem, gdyż dodatkowe ograniczenia czynią problem bardziej czytelnym, ale warto byłoby przedstawić również wersję uproszczoną, która może być łatwiejsza do optymalizacji.

Po drugie, tak sformułowany problem nie wymusza przypisania graczom preferowanych przez nich ról (co wcześniej postuluje autor). Mogą go minimalizować przypisania gdzie graczom przydzielane są nielubiane przez nich role o ile tylko sytuacja ta będzie miała miejsce w obu drużynach i będzie dotyczyła (średnio) wszystkich członków każdej drużyny. Problem ten dotyczy również proponowanego algorytmu.

Należy jednak zauważyć, że uwzględnienie umiejętności użytkowników z pewnością przynosi korzyści co widać wyraźnie w rezultatach eksperymentów.

W kolejnej części pracy autor przedstawia analizę zależności przestrzennych między członkami drużyny a wynikiem rozgrywki w grze League of Legends. W tym celu liczy szereg zmiennych związanych z przestrzennym rozkładem członków drużyny w kolejnych momentach czasowych. Dotychczasowe podejścia oparte o tę metodykę bazowały na małych danych, co autor postanowił naprawić. Doktorant wprowadził też dodatkowe korekty, np. słusznie usunął zmienną 'remoteness' jako mającą bezpośredni związek z wynikiem gry.

Na tak przygotowanych danych zbudowano las losowy a poprzez analizę jego istotności zmiennych zidentyfikowano czynniki wpływające na skuteczność gry danej drużyny. Przede wszystkim ważnym aspektem okazała się kontrola dużych fragmentów pola gry. Na tej podstawie autor przedstawił szereg rad zarówno dla graczy jak i dla programistów opracowujących zasady gry.

Ogólnie wyniki przedstawionych analiz są ciekawe a przedstawione rady powinny zainteresować graczy. Pewną wadą podejścia jest dość skromy zestaw zmiennych, które autor analizował. Ciekawe byłoby np. dodanie zmiennych związanych z czasem, np. szybkość poszerzania pola gry.

Trzecia część pracy dotyczy problemu modelowania funkcji wartości gier. W części tej doktorant stosował metody uczenia maszynowego do przybliżenia funkcji wartości na podstawie historycznych rozgrywek prowadzonych

przez boty. Zbiór danych był dostępny w ramach konkursu. Autor zastosował podejście oparte na łączeniu predykcji wielu modeli przy pomocy dodatkowego modelu predykcyjnego korzystającego z ich wyników. Stosował przy tym szereg nowoczesnych metod uczenia maszynowego takich jak LightGBM. Stosowane metody świadczą o dobrej znajomości uczenia maszynowego przez doktoranta.

Dodatkowo, w danych występował problem różnych rozkładów w zbiorze uczącym i testowym. Bardziej konkretnie, w zbiorze testowym występowały karty, które nie występowały w zbiorze uczącym. Autor rozwiązał ten problem adaptując metodę zaproponowaną wcześniej przez Sugiyamę [136]. Adaptacja polegała na modyfikacji procedury walidacji krzyżowej tak by w zbiorze testowym zawsze występowała określona liczba nowych kart. Metodę uznaję za twórcze i interesujące rozszerzenie istniejących metod.

Dzięki powyższym zabiegom autor uzyskał bardzo dobry wynik klasyfikacji, porównywalny ze zwycięzcą konkursu.

## Poprawki redakcyjne

Praca jest starannie opracowana pod względem redakcyjnym. Struktura pracy jest poprawna, bibliografia bardzo obszerna. Umieszczony na początku pracy słowniczek ułatwia lekturę tekstu osobie nieobeznanej z tematyką gier komputerowych.

Kilka literówek i drobnych błędów językowych wymieniam poniżej:

str. 16 ‘as far as in’ powinno być ‘as far back as’

str. 17 ‘threatening’ powinno być ‘treating’

str. 28 ‘nexu’ powinno być ‘nexus’

str. 30 podpis rysunku ‘Once’ powinno być ‘One’

str. 40 ‘enlisted’ powinno być ‘listed’

str. 53 ‘Thank to that’ powinno być ‘Thanks to that’

str. 55 ‘they trully exists’ powinno być ‘they trully exist’

str. 57 ‘onyl’ powinno być ‘only’

str. 67 ‘using wards make’ powinno być ‘using wards makes’

## Podsumowanie

Rozprawa odnosi się do trzech problemów związanych z grami komputerowymi i e-sportem: łączenia graczy w zespoły, analizy czasowej i przestrzennej strategii graczy oraz modelowania funkcji wartości. W każdą z tych dziedzin autor wniósł twórczy wkład. Pomimo pewnych niedociągnięć formalnych

w części pierwszej wkład ten należy uznać za wystarczający dla uzyskania stopnia doktora.

Podsumowując, stwierdzam że praca spełnia z wymagania ustawowe stawiane dysertacjom doktorskim w dziedzinie informatyki i stawiam wniosek o dopuszczenie doktoranta do publicznej obrony.

dr hab. inż. Szymon Jaroszewicz