

STATYSTYCZNA WERYFIKACJA „DLC” TEKKEN 7

Robert Woźniak  <https://orcid.org/0009-0009-6020-5247>

Wydział Zastosowań Informatyki i Matematyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: robo.wozniak@gmail.com

Streszczenie: Głównym celem pracy jest weryfikacja czy postacie dodatkowo płatne mają przewagę nad postaciami startowymi w grach akcji. Cel pracy zrealizowano przeprowadzając statystyczną analizę zwycięstw w grze TEKKEN 7 dla wszystkich 50 postaci (wojowników) występujących w grze. Do wyodrębnienia jednorodnych grup postaci wykorzystano metody *k*-średnich, grupowania hierarchicznego oraz model based clustering. Dodatkowo za pomocą metody TOPSIS został zbudowany ranking postaci.

Słowa kluczowe: TEKKEN 7, wielowymiarowa analiza statystyczna, metoda *k*-średnich, klastrowanie hierarchiczne, model based clustering, TOPSIS

JEL classification: C38, M210

SŁOWNIK POJĘĆ

1. „Fighting games”- bijatyka, rodzaj gier wideo skupiającej się na walce dwóch wojowników, którzy dysponują różnymi stylami walki.
2. „Data frame” – właściwości ruchu, każdy ruch ma indywidualne parametry, gdy uderzy, jest zablokowany, czy uniknięty w jakikolwiek inny sposób.
3. „Fighter” – wojownik, postać, charakter – dostępny w grze, istnieje podział na: „startowy” oraz „DLC”.
4. „DLC” – postać dodatkowa, dostępna odpłatnie. Ma na celu zwiększyć zainteresowanie grą, przedłużyć jej cykl życia oraz zwiększyć zysk z sprzedaży gry. Nie można jej (postaci) uruchomić bez zakupu gry.

<https://doi.org/10.22630/MIBE.2023.24.4.21>

5. Patch – aktualizacja usprawniająca działanie gry.
6. Kanibalizm – zjawisko, które występuje podczas wprowadzania gier do sprzedaży bardzo bliskim odstępem czasowym. Zainteresowany gracz, mimo, że chce zagrać w obie gry, musi wybrać tylko jedną ponieważ fizycznie nie jest w stanie grać w dwie gry.
7. Cykl życia gry – popularność gry na rynku, mierzona w godzinach spędzania graczy w grze.
8. „TIER LIST” – ranking postaci uszeregowanych od najlepszej do najgorszej. Postacie są umieszczone w grupach oznaczony często od liter „S” do „E”. Istnieje możliwość zmiany oznaczeń wedle twórcy.
9. Sklepy cyfrowe – usługi elektroniczne oferujące dystrybucje gier i rzeczy z nimi związanych oraz monitorowanie cyklu życia gry, śledzenie błędów oraz miejsce zamieszczania opinii graczy.
10. „VS”- versus, w badaniu będzie oznaczać liczbę wygranych meczy do wybranego przeciwnika.

Uwaga

Wymienione pojęcia nie są opracowaniem mechaniki gry. Są wprowadzeniem do artykułu w celu łatwiejszego zrozumienia przeprowadzonego badania, nie jest to także żaden przewodnik do gry ani poradnik do mechaniki.

WSTĘP

Od początku powstania „Fighting Games” gracze sondują, które postacie są najlepsze i czym wyróżniają się na tle innych. Użytkownicy gier sprawdzają wybrane cechy każdej z postaci takie jak: liczba wygranych turniejów przez daną postać, statystyki wygranych, łatwość obsługi postaci („Move set, Data frame”). Gracze dyskutują na temat każdej postaci a konsekwencją tych dyskusji są listy rankingowe - „TIERLIST”. Rankingi te służą do wstępnego rozeznania potencjalnego przeciwnika – postaci z którą gracz chce się zmierzyć. W tabelach 1. i 2. zaprezentowano przykładowe „TIER LIST”.

Tabela 1. Przykładowa TIER LIST TEKKEN 3

TIER	Postaci
S:	(True)Ogre
A:	Jin, Xiaoyu, Heihachi, Law
B:	Lei, Nina, Julia, Yoshimits, Anna, Mokujin
C:	Eddy, Paul, King, Hwoarang, Gun Jack
D:	Bryan, Kuma/Panda
E:	Dr. B, Gon

Źródło: opracowanie własne na podstawie danych zaczerpniętych z strony internetowej: wiki.supercombo.gg (dostęp z dnia 04.05.2022)

Tabela 2. Przykładowa TIER LIST TEKKEN 6

TIER	Postaci
S:	Nina Williams, Lee Chaolan, Alisa Bosconovitch, Armor King, Yoshimitsu, Devil Jin, Kazuya Mishima, Bryan Fury, Jin Kazama, Lili Rochefort
A:	Raven, King, Wang Jinrei, Zafina, Ling Xiaoyu, Jack 6, Paul Phoenix, Hwoarang, Sergei Dragunov, Heihachi Mishima, Marshall Law, Asuka Kazama
B:	Steve Fox, Miguel Caballero Rojo, Devil Kazuya, Baek Doo San, Eddy Gordo, Ganryu, Lei Wulong, Anna Williams, Leo Kliesen
C:	Lars Alexandersson, Julia Chang, Marshall Law, Feng Wei, Panda, Wang Jinrei, Roger Jr

Źródło: opracowanie własne na podstawie danych zaczerpniętych z strony internetowej: wiki.supercombo.gg (dostęp z dnia 04.05.2022)

Wraz z postępującym rozwojem technologicznym, powstawaniem baz danych, łatwiejszym dostępem do Internetu, bankowości elektronicznej, lepszych podzespołów komputerowych pojawiły się sklepy cyfrowe – platforma zarówno dla producentów jak i graczy. W tym miejscu producenci mogą dystrybuować swoje gry, monitorować cykle życia tych produktów oraz analizować reakcje społeczności graczy. Gracze z kolei mają ułatwiony dostęp do interesujących ich produktów, mogą oceniać gry oraz dyskutować na ich temat.

Po powstaniu sklepów cyfrowych przyszła moda na patch oraz promocję „DLC” w gatunku bijatyk. Bardzo często deweloperzy rozszerzają grę o dodatkowe postacie (DLC) ale są one sprzedawane po premierze gry. Wprowadzenie dodatkowych postaci spowodowało uaktualnienie TIER LIST. Gracze obserwujący

zmiany w rankingach analizowali dlaczego wprowadzenie jednej dodatkowej postaci zmienia układ rankingu oraz zaburza ustalony (poprzedni) porządek.

Głównym celem pracy jest weryfikacja, czy postacie dodatkowo płatne mają przewagę nad postaciami startowymi. Innymi słowy celem badania jest odpowiedź na następujące pytania:

1. Czy „DLC” jest dobrym zabiegiem marketingowym?
2. Czy „DLC” lepiej wypadają na tle postaci standardowych?
3. Czy można określić jakiegokolwiek powiązanie pomiędzy „DLC” ?

BADANIA LITERATUROWE

W wielu artykułach przeprowadzono analizy statystyczne w zakresie różnych gier komputerowych. W publikacji pod tytułem “Positive and Negative effects of Game Analytics in the Game Design” [Powell 2016] zbadano mocne i słabe strony budowy gier metodą “Grounded Theory”. W pracy pokazano, że za pomocą tej techniki można analizować zainteresowanie graczy wybranymi elementami gier, takimi jak dodatkowe darmowe dodatki („DLC”) czy nowymi darmowymi (dodatkowymi) poziomami gier.

Celem badania “Strategic Timing of Entry: Evidence from Video Game” [Engelstätter, Ward 2013] była statystyczna weryfikacja strategii wejścia produktu na rynek. Stwierdzono, że odpowiedni termin wprowadzenia gry na rynek minimalizuje efekt “kanibalizmu gier wideo”. Podobne informacje można odnaleźć w pracy “Statistical analysis of Game Behaviour” [Brodzki, Fleischer, Sarnik 2010]. Wnioski zaprezentowane potwierdzają występowanie “kanibalizmu” gier. Dodatkowo w tym artykule stwierdzono bardzo niską korelację pomiędzy recenzjami zawodowych recenzentów na stronach internetowych a cyklem życia gry. Przedmiotem badania zaprezentowanym w artykule “Positioning Game Review as a Crucial Element of Game” [Tong 2021] był wpływ recenzji użytkowników na sprzedaż gry. W konkluzji zaznaczono, że dłuższy okres życia gry nie wpływa na sprzedaż. Deweloperzy powinni skupić się na mechanizmach ułatwiających rozgrywkę nowym graczom, niż skupiać się na obecnych użytkownikach.

METODYKA BADAŃ

Metoda k-średnich:

Metoda k-średnich to technika klastrowania, która grupuje dane na podstawie ich podobieństwa. Głównym celem jest podzielenie zbioru danych na k klastrów, gdzie każdy klaster zawiera punkty danych podobne do siebie nawzajem. Dokładniejszy opis można znaleźć np. [Kassambara 2017].

Metoda grupowania hierarchicznego:

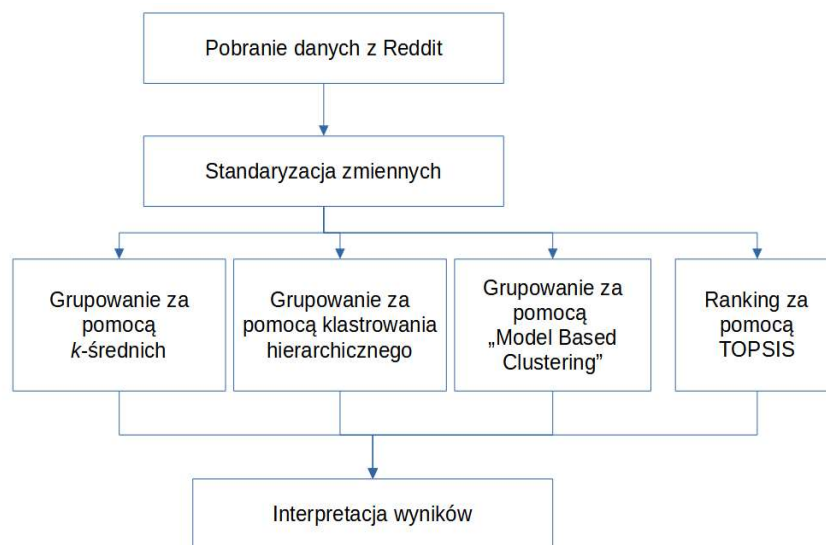
Do podziału analizowanego zbioru danych na grupy jest wykorzystywana metoda deglomeracyjna. Wynikiem hierarchicznego grupowania jest zestaw zagnieżdżonych klastrów, zwykle prezentowanych jako dendrogram - wielopoziomowe drzewo, w której klastry na jednym poziomie są połączone, tworząc większe klastry z jednego poziomu na drugi. Pozwala określić poziom, na jakim drzewo ma zostać przycięte, aby wygenerować odpowiednią liczbę klastrów. Dokładniejszy opis można znaleźć np. [Johnson, Wichern 2008].

Model based clustering:

Model based clustering to technika klastrowania, która opiera się na założeniu, że dane pochodzą z pewnego modelu statystycznego. W przeciwieństwie do tradycyjnych metod, takich jak k -średnich, w których klastry są identyfikowane na podstawie odległości między punktami danych, model-based clustering zakłada, że dane w klastrze podlegają pewnemu rozkładowi statystycznemu. Dokładniejszy opis można znaleźć np. [Bouveyron, Celeux, Murphy, Raftery 2019].

Szczegółowy schemat badania zaprezentowano na rysunku 1.

Rysunek 1. Poszczególne etapy prowadzenia badania.



Źródło: opracowanie własne

MATERIAŁ EMPIRYCZNY

Do badania wykorzystano dane ze strony Reddit [Reddit 2020], zawierające informacje typu „VS”. Przykładowy fragment danych zaprezentowano w tabeli 3.

Tabela 3. Dane wykorzystane do analizy empirycznej (fragment badania)

	VS vs. Akuma	VS vs. Alisa	VS vs. Anna	VS vs. Armor King
Akuma		47	48	46
Alisa	53		50	52
Anna dlc	52	50		53

Źródło: opracowanie własne na podstawie danych pochodzących ze strony: www.reddit.com

W grze TEKKEN 7 dostępnych jest 50 postaci w tym 11 „DLC”. Zgromadzone informacje z Reddita zawierają wyniki rywalizacji pomiędzy dwoma różnymi wojownikami. Przykładowo postać „Alisa” ma 50 zwycięstw przeciwko postaci „Anna”. Analogicznie się postępuje w przypadku innych wojowników. Puste pola zostały zastąpione najczęstszą występującą wartością w danej kolumnie, z powodu, że bezsensowne jest badanie zwycięstw tej samej postaci do niej samej.

WYNIKI BADAŃ EMPIRYCZNYCH

W tabeli 4 i tabeli 5 zaprezentowano odpowiednio wyniki uzyskane z metod klastrowania hierarchicznego oraz metody k -średnich. Za pomocą obu metod uzyskano tą samą liczbę klastrów, o bardzo podobnej strukturze. Innymi słowy, wyniki uzyskane dla obu metod są bardzo zbliżone – występują jedyne nieliczne przesunięcia pomiędzy grupami. W obydwu przypadkach można zauważyć odrębne grupy złożone z postaci dodatkowych (grupa 1. i 2.), grupę w której występuje silna dominacja „DLC” nad startowymi (grupa 3.) oraz pozostałych grup złożonych z podstawowych wojowników.

Tabela 4. Wyniki klastrowania hierarchicznego

Klaster	Nazwa Postaci
1	Leroy dlc, Ganryu dlc, Zafina dlc
2	Anna dlc, Kuma, Eddy, Jack-7, Eliza dlc, Panda, Gigas, Geese dlc, Marduk dlc, Lucky Chloe, Negan dlc, Claudio, Bob, Alisa
3	Noctis dlc, Paul Phoenix, Akuma, Armor King dlc, Katarina, Lili, Julia dlc, Nina, Lars
4	Feng, Heichachi, Shaheen, Hwoarong, Miguel, King, Bryan, Dargunov, Josie, Leo, Asuka, Lee, Master Raven, Lei dlc

Klaster	Nazwa Postaci
5	Xiaoyu Ling, Steve Fox, Yoshimitsu, Kazumi, Kazuya, Devil Jin, Jin

Źródło: opracowanie własne

Tabela 5. Wyniki klastrowania metodą k -średniej

Klaster	Postać
1	Leroy dlc, Ganryu dlc, Zafina dlc
2	Anna dlc, Kuma, Eddy, Jack-7, Eliza dlc, Panda, Gigas, Geese dlc, Marduk dlc
3	Noctis dlc, Paul Phoenix, Akuma, Armor King dlc, Katarina, Lili, Julia dlc, Nina, Lars, Feng, Heichachi, Shaheen, Hwoarong, Miguel, Josie, Lee, Master Raven, Lei dlc
4	Jin, Yoshimitsu, Steve Fox
5	Xiaoyu Ling, Kazumi, Kazuya, Devil Jin, Asuka, King, Bryan, Law, Leo, Dragunov

Źródło: opracowanie własne

W tabeli 6 zaprezentowano wyniki grupowania za pomocą modelu based clustering. Można zauważyć w niej mniej klastrów – dwa pierwsze mają wyraźną przewagę „DLC” a pozostałych głównie klastrach występują głównie postacie standardowe.

Tabela 6. Wyniki klastrowania model based clustering

Klaster	Postać
1	Fakhumram dlc, Kunimitsu dlc, Negan dlc, Alisa, Claudio, Marduk dlc, Leroy dlc, Gigas, Ganryu dlc, Anna dlc, Lucky Chloe, Zafina dlc, Panda, Kuma, Geese dlc
2	Eddy, Eliza dlc, Jack 7, Feng, Bob, Lars, Nina, Lili, Julia dlc, Katarina, Nina Paul Phoenix, Armor King dlc, Noctis dlc
3	Master Raven, Lei dlc, Heichachi Shaheen, Miguel, Bryan, King, Hwoarang, Josie, Dragunov
4	XiaYu Ling, Steve Fox, Asuka, Yoshimitsu, Devil Jin, Law, Kazumi, Kazuya, Jin, Akuma

Źródło: opracowanie własne

Ostatnią częścią badania była budowa rankingu za pomocą metody TOPSIS. Fragment rankingu zaprezentowano w tabeli 7. Pierwsze cztery miejsca zajmują postacie „DLC”. W częściowym zaprezentowanym rankingu można znaleźć 7 z 11 dostępnych płatnych postaci.

Tabela 7. Wyniki TOPSIS (wyniki częściowe)

Ranking	Nazwa postaci	Wynik	Ranking	Nazwa Postaci	Wynik
1	Leroy dlc	0.85745	8	Kuma	0.60897
2	Ganryu dlc	0.77017	9	Marduk dlc	0.58975
3	Zafina dlc	0.71553	10	Lucky Chloe	0.58838
4	Anna dlc	0.66565	11	Negan dlc	0.58505
5	Claudio	0.65896	12	Panda	0.57936
6	Alisa	0.65681	13	Gigas	0.57113
7	Geese dlc	0.61706	14	Eliza dlc	0.56353

Źródło: opracowanie własne

PODSUMOWANIE I WNIOSKI

W pracy zaprezentowano wyniki grupowania postaci gry TEKKEN 7 za pomocą trzech metod k -średnich, grupowania hierarchicznego oraz model based clustering. Za pomocą dwóch pierwszych uzyskano trzy skupienia zdominowane przez „DLC”. Dodatkowo można zauważyć wyraźny podział pomiędzy „DLC” a postaciami standardowymi. Wyniki uzyskane za pomocą model based clustering potwierdzają poprzedni podział. W rankingu zbudowanym metodą TOPSIS można zauważyć, że postacie „DLC” pod względem liczby zwycięstw są lepsze od postaci standardowych.

Po przeprowadzonych analizach można sformułować następujące odpowiedzi na pytania badawcze:

1. „DLC” jest dobrym zabiegiem marketingowym, odnoszą dużą liczbę zwycięstw, są pożądanymi postaciami dla graczy zatem generują dodatkowe zyski dla producenta.
2. Na podstawie przeprowadzonych analiz można zauważyć wyraźną przewagę „DLC” nad wojownikami standardowymi
3. „DLC” znajdują się w osobnych klastrach zatem można tylko przypuszczać o związku pomiędzy nimi. Jednak badanie nie jest wyjaśnieniem całkowitych

różnic pomiędzy indywidualnymi wojownikami, niezbędne jest także wykorzystanie innych metod, na przykład metod sieci, aby zanalizować przewagę na poziomie jednostkowym.

Podsumowując: za pomocą wykorzystanych metod uzyskano odpowiedzi na postawione pytania. Stwierdzono, że istnieje różnica pomiędzy „DLC” a postaciami standardowymi. Płatne postacie są lepsze pod względem liczby zwycięstw, będą więc chętniej wybierane przez graczy. W konsekwencji deweloper będzie rozwijał tą dodatkową opcję. Taki zabieg marketingowy potencjalnie przyniesie większe zyski.

Prezentowane badanie jest załącznikiem do dalszych badań w dziedzinie „Fighting games” oraz wyłonienia działań dewelopera jak i zachowań graczy.

BIBLIOGRAFIA

- Bouveyron C., Celeux G., Murphy T., Raftery A. (2019) *Model-Based Clustering and Classification for Data Science: With Applications in R (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press. doi:10.1017/9781108644181.
- Engelstätter B., Ward M. (2013) *Strategic Timing of Entry: Evidence from Video Games*. SSRN Electronic Journal, <https://doi.org/10.2139/ssrn.2379471>.
- Fleischer G., Sarnik K., Brodzki E. (2010) *Statistical Analysis of Gamer Behavior*. Worcester Polytechnic Institute.
- Johnson R. A., Wichern D. W. (2008) *Applied Multivariate Statistical Analysis*. New Jersey, Prentice Hall.
- Kassambara A. (2017) *Practical Guide to Cluster Analysis in R, Unsupervised Machine Learning*. STDH.
- Powell R. (2016) *Positive and Negative Effects of Game Analytics in the Game Design process: A Grounded Theory Study*. <https://www.semanticscholar.org/paper/Positive-and-Negative-effects-of-Game-Analytics-in-Powell/8fb33e472e2e03ce4a3f0aece75f222d68310183> [dostęp: 02.02.2023].
- Tong X. (2021) *Positioning Game Review as a Crucial Element of Game User Feedback in the Ongoing Development of Independent Video Games*. *Computers in Human Behavior Reports*, 3, 100077. <https://doi.org/10.1016/j.chbr.2021.100077>.

MULTIDIMENSIONAL STATISTICAL ANALYSIS TEKKEN 7

Abstract: The main goal of the study is to verify whether premium characters have an advantage over the starter characters. The objective was achieved by conducting a statistical analysis of victories in the TEKKEN 7 game for all 50 characters (warriors) present in the game. Methods such as *k*-means clustering, hierarchical clustering, and Model-Based Clustering were employed to identify homogeneous groups of characters. Additionally, a ranking based on the applied methods was constructed using the TOPSIS method.

Keywords: *k*-means, hierarchical clustering, model based clustering, TOPSIS

JEL classification: C38, M210