

Csertán Tamás

Emberi navigáció vizsgálata hálózati játékok segítségével

Szakmailag ellenőrizte: Dr. Gulyás András

(egyetemi docens, BME Távközlési és Médiainformaticai Tanszék)

A hálózatelmélet szinte gyermekkorúnak számító tudományág, melyben számos jelenségnek még nincs kiforrott magyarázata. Ezek közé tartozik a navigáció komplex hálózatokban, ahol a korábban uralkodó 'legrövidebb utakat használjuk' felfogást utóbb több kutatás árnyalta. Nem szükséges hangsúlyozni, hogy mennyire fontos erről valódi tudást szerezni: kognitív és valós tereinkben való útkeresésünk, az adatcsomagok útja az interneten és az ingerületeké az agyban mind lefordíthatóak hálózati navigációs feladatok megoldására. Dolgozatomban áttekintem a hálózati navigációra felállított különféle elméleteket és ismertetem a legfrissebb kutatási fejleményeket. Bemutatom egy Androidos játék elkészítését, mely komplex hálózaton alapul, s segítségével vizsgálható játékosok navigációs viselkedése. Leírom hipotéziseimet az emberi navigációs stratégiákról, melyeket a játék eredményei, a Wikipediagame adatbázisban fellelhető adatok és mai tudományos kutatások alapján fogalmazok meg, majd megvizsgálom érvényesülésüket ezekben az adathalmazokban.

1. Bevezetés

Mikor a világ jelenségeit szemléljük, a felszínen látható hatalmas változatosságra tekintve nem is gondolnánk, hogy mindeme színes jelenségek rendkívül széles skáláját rejtetten ugyanazok a belső struktúrák, szabályok és dinamikai tulajdonságok határozzák meg. Lehetséges, hogy első pillantásra nem fedezzük fel a hasonlóságot a hazai vízi ökoszisztéma és az internet, a nemzetközi tudományos publikációs rendszer és az idegsejtek hálózatai között, ám a XX. század végén kialakult hálózattudomány sorozatos felismerései ezeket a színleg távol álló területeket mára szoros közelségbe hozták egymáshoz (Barabási & Pósfai, 2016). Világunkat át- és áthatják a hálózatok, és ennek az univerzalitásnak tudatosításával a hálózattudomány szerepe és fontossága is töretlenül növekszik a tudományos közösségben.

Noha manapság egyre több kutató kapcsolódik be a tudományterületen folyó munkába, és az empirikus vizsgálatok elvégzéséhez is egyre több adat áll rendelkezésre a különböző adatrögzítési és feldolgozási technológiák fejlődésének köszönhetően, egyes jelenségekre és problémakörökre még nem állnak készen teoretikus magyarázatok és minden aspektusukat kielégítően megmagyarázó elméletek.

Ilyen terület a hálózatokon belüli navigáció kérdése is (Zhu & Levinson, 2015). Mikor üzeneteket, információt, jeleket vagy akár fizikai objektumokat szeretnénk eljuttatni egyik helyről a másikra akár az interneten, akár fizikai valóságunkban, valójában egy hálózatban kell megoldanunk egy navigációs feladatot. Természetesen bizonyos feltételeknek meg kell

felelnünk, melyek saját magunk által szabottak vagy a környezet által előírtak: hatékony megoldásra kell jutnunk az eltelt idő, ráfordított energia, igénybe vett számítási kapacitás és még rengeteg lehetséges jellemző szempontjából; emellett esetleg előnyben kell részesítenünk egyes útvonalakat és csomópontokat, míg másokat mellőznünk szükséges, avagy ki kell egyensúlyoznunk az egyes helyekre eső forgalmi terhelést. Összefoglalva: útkeresésünk során óhatatlanul optimalizálunk.

Hálózatelméleti kutatások során a kutatók egyszerűsítésként előszeretettel tételezik fel, hogy a rendszerben található ágensek az abszolút optimális, legrövidebb utakat találják meg és használják ezeknek a feladatoknak a megoldása során (Newman, 2010). Azonban az utóbbi időben több tanulmány is rámutatott arra, hogy a valóságban használt effektív útvonalak megnyúlnak a feltételezettekhez képest, azaz hosszuk nagyobb lesz, mint a legrövidebb utaké (Csoma, Kőrösi & Rétvári, 2017). Erre a jelenségre csupán józan belátás alapján is találhatóak már viszonylag szemléletes magyarázatok. Ha értelemmel nem rendelkező ágenseket tekintünk, ahol példának mondjuk az internet egyes csomópontjaiban lévő routereket vesszük, a feladat egyensúlyi tényezői a következőképpen alakulnak: minden router optimális utat tudna választani az adatcsomagjainak, ha ismerné a hálózat teljes topológiáját – ezt a hálózat méretét tekintve nem racionális feltételezni –, ugyanakkor minél kevesebb lokális információval rendelkezik erről a szerkezeetről, annál pontatlanabban tudja kézbesíteni azokat. Tehát az információk teljessége és az útvonalaink optimalitása szoros kapcsolatban áll egymással.

Dolgozatomban olyan feltevéseket vizsgálok, amelyek képesek tisztábbá tenni a képet a komplex hálózatokban végzett navigáció egyes tulajdonságait tekintve. Szűkebb körben, az emberek által követett stratégiákat vizsgálom meg két kísérleten keresztül, melyeknek eredményei tágabb kontextusukban is értelmezhetők, mivel az emberi elme viselkedését e szűkített esetben is a teljes informáltság és ideális navigáció közti dilemma karakterizálja csakúgy, mint az általános esetben.

Megvizsgálom a bejárt utak megnyúlását és egy egyszerűsített tájékozási törzshálózat kialakulását a *The Wiki Game* (Clemesha, 2020) című online játék felhasználóiról készült adatbázisban, mely a tartalmazott adatok sokszínűségére és számosságára tekintettel erős bizonyítékként szolgálhat a feltevések ellenőrzésénél. Majd bemutatom egy saját magam által tervezett, *Matrice* (Csertán, 2020c) nevű Androidos játék specifikálását, megvalósítását és a belőle nyert adatok elemzését, ami további ismereteket szolgáltat a törzshálózatok kialakulásának feltételeiről, illetve felveti az emberi stratégiák egyéb típusainak létezését is.

A második fejezetben rövid áttekintést adok a kutatásomhoz szorosan kapcsolódó fogalmakról, a hálózatok leírására használt eszköztár egyes elemeiről. Ezután a harmadik szekcióban ismertetem az általam tervezett játék tervezési és megvalósítási folyamatát, és az adatgyűjtést. Majd a záró fejezetben leírom hipotéziseimet és vizsgálatom eredményeit a két játékból nyert adatok elemzése alapján.

2. Hálózatelméleti alapok

A hálózatelmélet alapvetően empirikus tudományág, mely különböző forrásokból összegyűjtött nagy mennyiségű adatot elemez, alkot modellt és von le következtetéseket e modellek hasonlóságai és különbségei alapján. Elemzési eszköztáraként gráfelméleti fogalmakat használ, lévén a hálózatok remekül reprezentálhatóak a diszkrét matematikának ezzel a kiforrott eszközével. Az alapvető, gráfokra általánosan bevezetett és alkalmazott jellemzők (mint például a *fokszám*) mellett nagyobb ívű tulajdonságokat is leír (lásd *kisvilágosság*), melyek a természetben előforduló hálózatok körére egyöntetűen jellemzőnek mutatkoznak.

2.1 Komplex hálózatok általános vonásai

Bonyolult rendszerek sokasága vesz körül bennünket. A hálózatok kutatása a XX. század utolsó éveitől kezdte el felfedezni, hogy mind e rendszerek mélyén hálózatok rejtőznek, és hogy tulajdonságaikban teljesen eltérő előfordulási területük, méretük, természetük, koruk és formájuk ellenére nagyon sokban hasonlóak egymáshoz, és azonos rendezőelvek szabályozzák őket. Köszönhetően interdiszciplináris, kvantitatív és empirikus jellegének, a tudományág egyszerre próbál fényt deríteni e hálózatok statikus és dinamikus tulajdonságaira, foglalkozik időbeli fejlődésükkel és különböző zavarokkal szemben kifejtett ellenállóképességükkel.

1. ábra: Különböző valós hálózatok statikus tulajdonságai. Minden hálózathoz adott a csomópontok száma (Size), az átlagos fokszám ($\langle k \rangle$), az átlagos úthossz (ℓ) és a klaszterezettség (C). Összehasonlításképp szerepel egy azonos méretű véletlen gráf átlagos úthossza (ℓ_{rand}) és klaszterezettsége (C_{rand})

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	Reference	Nr.
WWW, site level, undir.	153, 127	35.21	3.1	3.35	0.1078	0.00023	Adamic 1999	1
Internet, domain level	3015 - 6209	3.52 - 4.11	3.7 - 3.76	6.36 - 6.18	0.18 - 0.3	0.001	Yook <i>et al.</i> 2001a, Pastor-Satorras <i>et al.</i> 2001	2
Movie actors	225, 226	61	3.65	2.99	0.79	0.00027	Watts, Strogatz 1998	3
LANL coauthorship	52, 909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman 2001a,b	4
MEDLINE coauthorship	1, 520, 251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman 2001a,b	5
SPIRES coauthorship	56, 627	173	4.0	2.12	0.726	0.003	Newman 2001a,b,c	6
NCSTRL coauthorship	11, 994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman 2001a,b	7
Math coauthorship	70, 975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási <i>et al.</i> 2001	8
Neurosci. coauthorship	209, 293	11.5	6	5.01	0.76	5.5×10^{-5}	Barabási <i>et al.</i> 2001	9
<i>E. coli</i> , substrate graph	282	7.35	2.9	3.04	0.32	0.026	Wagner, Fell 2000	10
<i>E. coli</i> , reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner, Fell 2000	11
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya, Solé 2000	12
Silwood park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya, Solé 2000	13
Words, cooccurrence	460,902	70.13	2.67	3.03	0.437	0.0001	Cancho, Solé 2001	14
Words, synonyms	22, 311	13.48	4.5	3.84	0.7	0.0006	Yook <i>et al.</i> 2001	15
Power grid	4, 941	2.67	18.7	12.4	0.08	0.005	Watts, Strogatz 1998	16
<i>C. Elegans</i>	282	14	2.65	2.25	0.28	0.05	Watts, Strogatz 1998	17

Forrás: Albert & Barabasi, 2001

Az **1. ábrán** néhány, a korai kutatásokban felmért hálózat tulajdonságai láthatóak, köztük az internet tartalmára (webes dokumentumok) és szerkezetére (fizikai struktúra), filmszínészek és tudósok kapcsolataira, természetes táplálékláncokra és az elektromos rendszerre vonatkozó példákkal. Egy pillantást vetve rájuk, rögvést észrevehetünk bizonyos hasonlóságokat: méretüktől függetlenül e hálózatok mindegyikének csúcsai átlagosan kevés kapcsolattal rendelkeznek, így maguk a hálózatok is ritkának mondhatóak. Mégis átlagos úthosszuk meglehetősen kicsiny nagyságukhoz képest. Lennie kell tehát egy különleges elrendezésnek, mely a hálózat e nagyfokú átjárhatóságát és relatív kis átmérőjét biztosítja. Ugyan a fenti adatok alapján az egyes csúcsokra és ezáltal a szerkezet hierarchikusságára vonatkozóan nem élhetünk alapos feltételezésekkel, annyit módunkban áll megjegyezni, hogy a véletlen gráf megfelelőjükhöz képest klaszterezettségük kifejezetten magas, azaz lokális szinten erős összefüggés áll fenn a csomópontjaik között.

A hálózatok tulajdonságai közül a legfontosabbak tehát, melyek vizsgálódásaim szempontjából kiemelkedően relevánsak: a skálafüggetlenség és a kisvilágosság.

2.2 Navigáció komplex hálózatokban

Bizonyos tulajdonságokkal rendelkező valóságos hálózatok hatékony módot kínálnak a bennük való tájékozódásra. A kis világokban gyorsan terjedhet az információ, skálafüggetlen hálózatok pedig optimális környezetet biztosíthatnak lokális tudás alapú navigációhoz. Milgram levélküldős kísérlete (Milgram, 1969) két roppant érdekes aspektusát mutatta meg a hálózatoknak: az elsőt, miszerint hatalmas méretük ellenére viszonylagosan rövid utak léteznek bennük a strukturális modellek jól megmagyarázzák; viszont az, hogy az emberek hiányos információkkal is hatékonyan találnak bennük utakat, máig nem értjük teljesen.

2.3 Emberi navigáció vizsgálata

A legtöbb tudományágban az elért elméleti eredmények precíz és alapos empirikus bizonyítékokkal való alátámasztására van szükség. Nincs ez másként a hálózattudomány és a benne felállított navigációs modellek esetében sem. Azonban az utóbbi időkig nem nagyon álltak rendelkezésre olyan adatok, amelyek a hálózatokban navigáló emberek valós útjait írták volna le (Csoma, Kőrösi & Rétvári, 2017). Emberek, adatok vagy jelek komplex rendszerekben megtett útjait megfigyelni egyébként sem triviális feladat – gondoljunk csak például a magas absztrakciós szinten lévő hálózatokra, mint a kognitív sémák, avagy a nyelv; vagy a megfigyelés magas költségeire: követőkódok számításgénye az interneten, adatvédelmi megfontolások az okostelefonos térkép appok esetében; és az értékelhető mennyiségű adat begyűjtésének fáradságosságára.

Megoldást jelenthetnek a felmerült nehézségekre a *hálózati játékok*. Dolgozatomban azokat a játékokat definiálom hálózati játéknak, amelyek elsődleges vagy mögöttes logikája reprezentálható egy gráffal, és a játéktevékenység egy része vagy egésze lefordítható a gráfon véghezvitt tájékozódási problémára. Ha jobban megvizsgáljuk őket, játékok tucatjai esnek – akár tervezői szándék nélkül – ebbe a kategóriába, több közülük világméretű hírnévre tett szert. Néhány példa a *PokemonGo*, a *WikiSpeedia* vagy a *fit-fat-cat* (két utóbbit tudományos kutatásra is felhasználták (West, 2009; Gulyás, 2020)). Játékos formában nem csupán az adatgyűjtés lehet teljes mértékben kontrollált és anonim, hanem szórakoztató és lekötő jellege megfelelő mennyiségű adattal láthatja el a jelenségek megértését hajszolókat.

3. Hálózati játék készítése

Saját játékom fejlesztésénél a következő szempontokat tartottam szem előtt: elsőként meg kell felelnie az imént kimondott definíciónak, hogy az emberi navigáció hálózatelmélet alapú vizsgálatára alkalmas legyen. Másodsorban alkalmas és élvezhető játéknak kell lennie, hogy a játékosokat megragadja és lekösse, ezáltal játék voltának is megfeleljen. Harmadrészt pedig szem előtt kell tartani egyszerűségét, mind a játéklógika, mind a megvalósíthatóság és a mögöttes hálózat leírhatóságának és elemezhetőségének szempontjából.

3.1 Tervezés

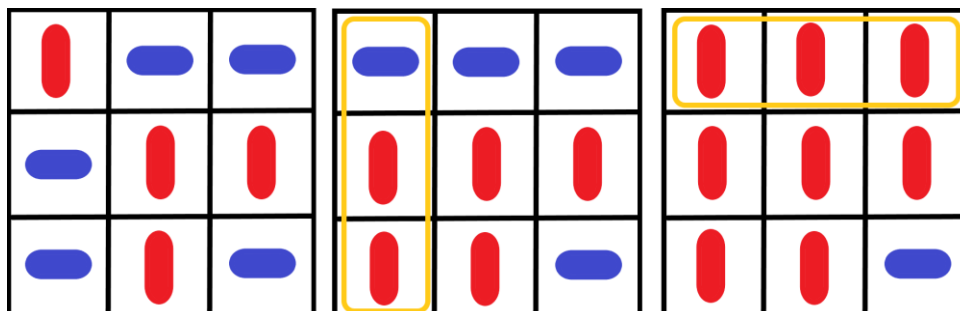
A tervezés első fázisában két példából merítettem inspirációt, a Wikispeedia és a fit-fat-cat nevű játékokból. Mindkét játék megvalósítani látszik a célkitűzéseimet, ugyanakkor azokból is hasznos volt őket jobban megvizsgálni, mivel hozzájuk képest merőben másfajta játékközpontú szerkezetet szerettem volna megvalósítani. A Wikispeedia játékot a Wikipédián is kimutatott kisvilág tulajdonság és az internetes szótár széleskörű hírneve tette népszerűvé. Játékosainak feladata, hogy egy-egy véletlenszerűen sorsolt kiindulási cikkből a cél cikkhez jussanak el, a lehető legkevesebb számú szócikk közbeiktatásával (és szócikktől szócikkig egyedül az oldalakon lévő hivatkozásokat követve szabad eljutni. A véletlen választás persze olyan mulatságos utakhoz vezet, mint amelyekben *Esterházy Péter* és a *guancsok* – egy kanári-szigeteki törzs – közötti kapcsolatot kell felkutatni¹.

Szavakat kell összekötniük a fit-fat-cat alkalmazással játszóknak: kezdetben adott egy hárombetűs angol szó (fit) és csupa értelmes szón keresztül, minden lépésben legfeljebb egy betűt megváltoztatva (fat) el kell érni egy végsőként megadott szót (cat). Ebben a játékban a szavak egy részhálózata a mögöttes gráf, éllel mindazon pontok között, melyek egyetlen betű megváltoztatásával egymásból megkaphatóak. Mindezek alapján az alábbi koncepciót vázoltam föl a saját játékomnak:

Az ötlet a számozás biciklilakatok ihletésére született, és leginkább logikai játékokra hajaz: adott egy feladvány, amelyet meg kell oldani a lehető legrövidebb idő és legkisebb lépésszám alatt; ehhez egy kezdeti állapotból egyszerű, elemi lépések segítségével kell eljutnia a játékosnak egy végső állapotba, ami a megoldást jelenti. Mivel így minden egyes lépés egy-egy újabb állapothoz vezet, és az állapotter kialakításában rögzített mennyiségű állapotot tartalmaz, a feladat megoldása megfeleltethető egy olyan gráfban való legrövidebb út keresésének, amelyben minden állapotot egy csúcshoz, és a köztük lévő átmeneteket egy-egy él jelenti meg.

A játék konkrét tervének elkészítése során próbáltam a valóban egyszerű logikai feladványok stílusát és nehézségét reprodukálni, ugyanakkor akkora állapotteret létrehozni, ami a benne való tájékozódást nem teszi sem triviálissá, sem elvégezzhetetlenül nehézé. Következésképpen néz ki a végleges feladvány: adott egy háromszor három mezőből álló pálya. Minden mezőben egy-egy figura helyezkedik el, ezeknek a figuráknak két állapota van, amint az a **3. ábrán** megfigyelhető.

3. ábra: Játéktér és egy hármas láncból álló mozgássorozat



Forrás: saját munka

¹ A legrövidebb út ebben az esetben háromlépéses: Esterházy Péter – 1950 – Berber naptár – Guancsok

Kétféle elemi lépést tehet meg a játékos: egy sorban vagy oszlopban cseréli meg a figurák állapotát az ellenkező állapotra (inverzió) vagy egy sorban vagy oszlopban elforgatja eggyel az ott található figurákat². Fent egy két lépésből – mégpedig két inverzióból – álló állapotsor látszik ennek megfelelően. A játék célja, hogy a játékos a lehető legrövidebb idő és lépésszám alatt egy sorsolt kezdőállapotból a szintén véletlenszerű végállapotba érjen. Ezalatt a mögöttes hálózat két pontja (a két állapot) között keres utat, úgy, hogy a definiált éleken (az egyes lépések) keresztül halad a célja felé. E módon a játék jól modellez egy navigációs helyzetet, ráadásul olyat, melyben az útkereső nem teljes információs állapotban van, hiszen elvárhatóan nem fogja tudni megjegyezni vagy kiszámítani az összes lehetséges következő állapotot, s ezáltal felderíteni a teljes gráfot, hanem legfeljebb az aktuális csúcs közvetlen szomszédairól deríthet fel bizonyos információkat.

Azért, hogy a játék mögöttes hálózatáról még több információt szerezzek, egy *Python* nyelvű programmal felépítettem a specifikációnak megfelelő gráfot, majd elemeztem annak tulajdonságait az *igraph* programcsomagot használva, ami gráfok analíziséhez szolgáltatt roppant kézre eső eszközöket (Csardi & Nepusz, 2006).

Az **1. táblázat** tartalmazza a részletes jellemzőket, melyek alapján a hálózat kisvilág tulajdonságot mutat, vagyis méretéhez (512 csúcs) képest viszonylag rövid utak találhatóak benne (három lépés körül átlagosan). Megfigyelhető továbbá az **1. ábra** adataival összevetve, hogy nagymértékben hasonlít az *E. coli* baktérium jellemzőire épített és néhány, táplálékláncok által meghatározott hálózatra, azzal a megszorítással, hogy klaszterezettségének mértéke – bár a véletlen gráfokéval nem esik egy nagyságrendbe – kisebb, mint a topológiaiilag szorosan korrelált hálózatoké, és a villamosenergia-hálózattal mutat leginkább hasonlóságot.

1. táblázat: A Matrice játék mögöttes hálózatának tulajdonságai

Tulajdonság	Érték
Csúcsok száma (N)	512
Élek száma (E)	5376
Átlagos fokszám ($\langle k \rangle$) ³	21
Átlagos úthossz (L)	3,18
Átmérő (D)	5
Klaszterezettség (C)	0,042

Forrás: saját munka

² Forgatás alatt azt a mozgást értjük, ami leírható a következőképp: legyen a sor tartalma x, x, y , ekkor ennek eggyel elforgatottja az x, y, x ; kettővel elforgatottja az y, x, x ; hárommal elforgatottja pedig önmaga.

³ Ki- és be fokszám egybevéve. Külön-külön mindkettő értéke 10,5.

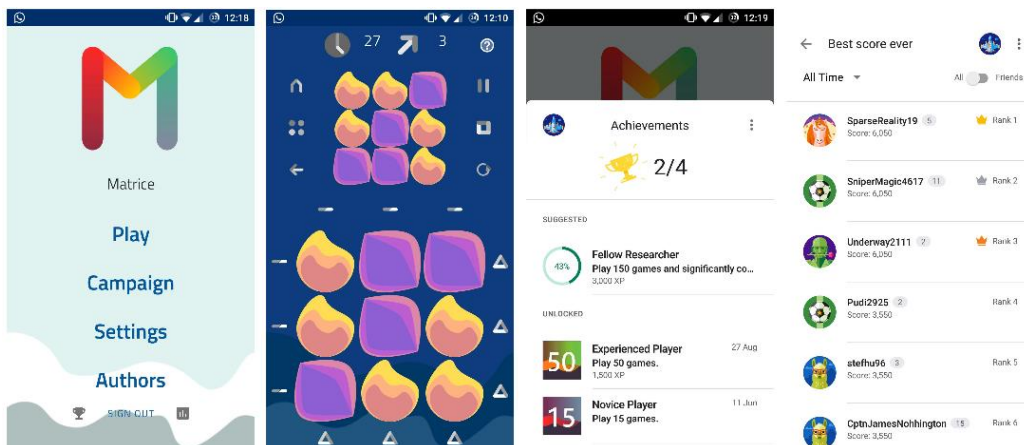
Leszűrhető a feltevés, hogy az emberek képesek lesznek rövid utakat találni a hálózatban, lévén az kisvilág tulajdonságú és kalkulálható logikájú, viszont valószínűleg eltérő stratégiát fognak alkalmazni, mint a Milgram-kísérlet alanyai, ahol a küldők segítségére volt a társadalom erős földrajzi tagoltsága; vagy a fit-fat-cat játékhoz képest, melyben a hálózat pontjaihoz erős szemantikus jelentés társult. Esetünkben más fogódzók vannak jelen, logikai szabályok és grafikus állapotmegjelenések, és az azokból szintetizálható vizuális mozgásformák.

Megterveztem magát az alkalmazást, tekintetbe véve mind a felhasználói élményt, mind a fejlesztői fenntarthatóságot. Figyelembe kellett venni, hogy a játéknak egyedi kinézetűnek és megragadónak kell lennie ahhoz, hogy a játékost lekösse, mivel az alapesetben 512 pontú hálózatban több tucat lejátszott játékra van szükség ahhoz, hogy az adatok elemzésénél konfidens eredményeket kaphassunk⁴. A játék emellett vizuális megjelenésében legyen egyszerű és egyedi. Ehhez a két kritériumhoz végig ragaszkodtam a tervezés során.

3.2 Megvalósítás

A játékot *Android* operációs rendszert futtató mobiltelefonokra készítettem el, hisz ezek elterjedtségükben és elért korosztályukban ideálisak egy játék kiadása szempontjából.

4. ábra: Az elkészült Matrice játék kinézete és néhány játékfunkció



Forrás: saját munka

A 4. ábrán látható az elkészült játék néhány minta-képernyője: a bejelentkezés után elérhető menü; maga a játéktér felső részén a cél állapottal, körülötte a funkciógombokkal, s alul az aktuális állapotot mutató játéktáblával; a játékelményt bővíteni hivatott, a *Google Play Games Services* segítségével elkészített küldetések és rangsorok. A kódolás során végig szem előtt tartottam azt a lehetőséget, hogy a munkámat esetleg más programozó/diák folytatni vagy kiegészíteni kívánja, ezért részletes dokumentációval láttam el a programkódot a *javadoc* formátumnak megfelelően, valamint publikusan elérhetővé tettem egy *Github repository*-n (Csertán, 2020a, 2020b).

⁴ Becslés a hivatkozott tanulmány alapján (Gulyas, 2020).

4. Eredmények

Az alkalmazást a fejlesztés és 8 ember bevonásával elvégzett teljeskörű bétatesztelés után 2020. 09. 08-án publikáltam a *Play* áruházban. Visszajelzések alapján az első publikus verziót különböző javításokkal még kétszer egészítettem ki, így nyerte el mai formáját. Meg kell jegyezni, hogy a játék jelen állapotában – bár a specifikált hármas cél (feladvány, megoldás, adatgyűjtés) megvalósítására tökéletesen alkalmas, nem tartalmaz számos olyan élvezetnövelő funkciót, amely a játékelményt valóban megragadóvá tenné (küldetés mód, visszajelzés a felhasználó által adott megoldás jószágáról, etc.), ezek megvalósítását idő hiányában halasztottam későbbre. A november 20-ig tartó adatgyűjtési periódus alatt 85 felhasználó regisztrált a játékban (ennél valamivel többen telepítették), ők összesen körülbelül 3800 játékot játszottak le. További statisztikák és az adatok részletekbe menő elemzése a Matrice adatbázis elemzése szakaszban találhatóak meg.

4.1 Hipotézis az emberi navigációra

Az utóbbi időkben több kutatás vizsgálta az emberi navigáció mögött álló tényezőket. Ennek motivációját részben a Milgram-kísérlet meglepő következtetése – hogy az ember képes rövid utakat találni komplex hálózatokban (Kleinberg, 2000) –, részben pedig az útvonalak utóbb megfigyelt megnyúlása – hogy közelítjük a legrövidebb utakat, de legtöbbször nem találunk optimális megoldást (Zhu & Levinson, 2015) – adta.

Megfogalmazódott, hogy az emberi navigációt bizonyos hálózati tulajdonságok segítik elő (Newman, 2003), hogy az alapvető stratégia a *'keress egy szupercsúcsot⁵, majd ereszkedj hazáig'* elvet követi (Leskovec, 2012), és hogy az utak választása során követett szabályok legerősebben a hálózatok hierarchikus tulajdonságait veszik figyelembe, semmint a legrövidebb utak iránti igényt (Csoma, Kőrösi & Rétvári, 2017).

Egy, a fit-fat-cat játékot vizsgáló tanulmány komplett magyarázatot javasol és bizonyít a szóhálózat gráfján, amit a következőkben én is analízisem tárgyává teszek.

A játékosok útkeresése során tapasztalható karakterisztikus nyúlás két szempont közti optimalizáció eredménye: az emberi elme egyensúlyt teremt a navigációjához szükséges helyben⁶ rendelkezésre álló információk és az útvonalak rövidege között. Mindehhez a játék használatának kezdetén egy tanulási periódus járul, melynek során alapvető információkat sajátít el a szóhálózatról és felépít egy sematikus, egyszerűsített törzshálózatot, ami az eredeti, meglehetősen terjedelmes gráf részgráfja. A törzshálózat később a navigációt nagyban meghatározza: egy-egy véltelenszerű csúcsból az útkeresési törekvés a törzshálózatba vezet, ahol a már ismert és begyakorolt utakon keresztül ér a cél közelébe.

Elmondható az is, hogy ez a törzshálózat minden emberre nézve egyedi, bár az egyes csúcsok beválogatódásának esélye a csúcsok hálózati tulajdonságaival korrelál (hiszen érdemesebb egy sok kapcsolattal rendelkező csúcsot bevenni a törzshálózatba). Összességében ez a leegyszerűsítő stratégia a szóhálózatban egy nagyságrendnyivel kevesebb értesülés⁷ tárolását teszi szükségessé a teljes hálózathoz képest, miközben az útvonalak nyúlása bőven egy elviselhető szint alatt marad (Gulyás, 2020).

⁵ olyan csúcs, amely kiemelkedően sok kapcsolattal rendelkezik, vagy különböző centralitás mértékek szerint a hálózat középponti részén helyezkedik el

⁶ a gráf egy adott lokációjáról és annak környezetéről

⁷ az információt a kutatásban az entrópia fogalmával mérték

Feltevésém szerint hasonló stratégia és olyasfajta törzshálózat kialakulása figyelhető meg a Wikipédián való navigációkor is, mint amelyet a kutatás leír. Vizsgálatom kiterjed a Matrice hálózatra is, amely gyökeresen más navigációs környezetet jelenít meg, és várakozásaim szerint megszorításokkal lép fel ezen emberi stratégia megvalósulásának feltételeit illetően, de legalábbis árnyalja azokat.

4.2 A hipotézis vizsgálata

A következőkben részletesen megvizsgálom azokat a feltételeket és körülményeket, amelyek a navigációt elősegítik, vagy éppenséggel lehetővé teszik a komplex hálózatokban. Mindehhez két hálózatból származó információkat elemzek: a Wikipédia hatalmas méretű, önszerveződő hálózat, mely természetes hálózatevolúciós folyamatok hosszú sorával jött létre; míg a Matrice játék mögött álló logikai struktúra méretben sokkal kisebb, mesterséges és merőben más tulajdonságokkal rendelkezik, mint nagyobb társa. A felhasználói adatokból leszűrt eredmények függvényében rá fogok mutatni ezen eltérések jelentőségére.

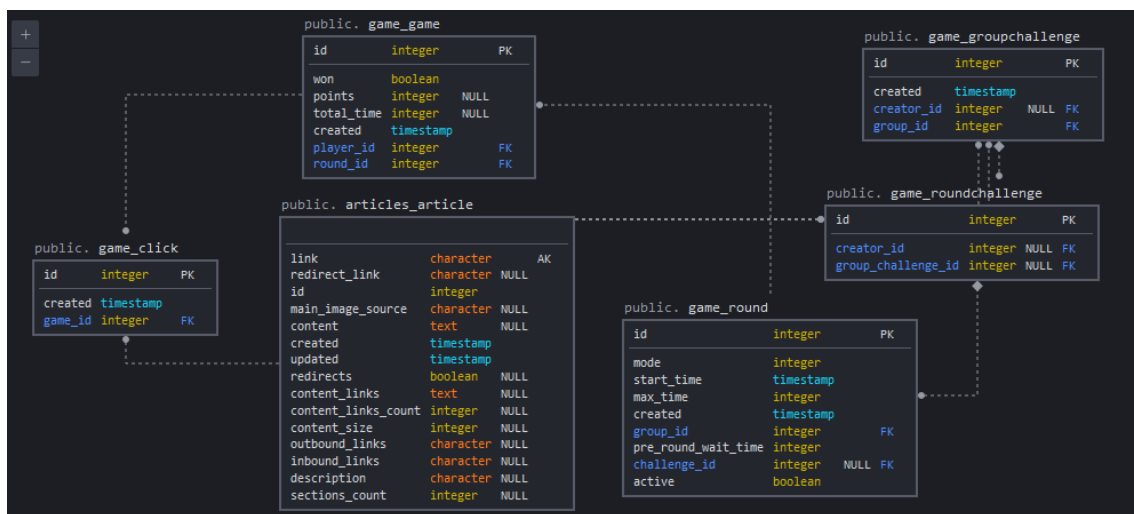
Dolgozatomban az eredmények bemutatását és a következtetések levonását helyezem fókuszba, így az egyes részszámítások eredményei és az elkészítésükhöz szükséges programkódok a szövegben táblázat-formátumban, valamint egy Github tárhelyen találhatóak meg (Csertán, 2020a). Az adatelemzési módszerek velős bemutatására természetesen a szövegben is sor kerül.

4.2.1 A The Wiki Game adatbázis elemzése – adatok és szerkezet

A The Wiki Game nevű játék a korábban bemutatott, Wikipédián játszható útkereső játék egyik implementációja, amit egy amerikai szoftverfejlesztő, *Alex Clemesha* készített el. A mögötte álló adatbázis a rendelkezésemre állt a dolgozatom készítésekor. A benne lévő adatok logikája meglepően egyszerű volt: adott egy kiindulási és egy érkezési cikk, rögzítendő az összes kattintás, amit a felhasználó végez.

A **5. ábra** bemutatja a PostgreSQL adatbázis szerkezetét. Megfigyelhető, hogy a játék, amiről egy külön táblában statisztikus információkat tárolunk kattintásokból áll össze, melyekből a rögzített időbélyegek alapján tudjuk a játékos játékmenetét rekonstruálni.

5. ábra: A The Wiki Game adatbázis PostgreSQL sémája



Forrás: saját fotó

Minden kattintás egy-egy cikkekre mutat, ami megfeleltethető a Wikipédia egy konkrét oldalának. Az egyes küldetések tartalmazzák a megoldandó feladatot (azaz a kiindulási és érkezési cikkpárt), ezeket több játékos is megpróbálhatja megoldani, ezzel létrehozva a konkrét játékokat. Szinte teljesen megegyezik ezzel a szerkezettel a MySQL adatbázis struktúrája is, eltekintve egyes apró adattárolási különbségektől és bizonyos mezőknek a táblákban elfoglalt helyétől.

Összességében a két adathalmazrész több mint 15,6 millió játék adatait tartalmazza, több mint 780.000 játékostól. Ez átlagban 20 játékot jelentene játékosonként, ami a játszmák alacsony száma miatt nem adna lehetőséget a hipotézis vizsgálatához, lévén maga a Wikipédia ennél nagyságrendekkel nagyobb gráf; ám roppant előnyös módon a játékosok között előfordulnak olyanok, akik megfelelő mennyiségű menetet lejátszottak a sikeres elemzéshez.

Vizsgálataim a 262 legtöbb játékkal rendelkező személyt érintették. A **2. táblázat**ban gyűjtöttem össze egyes imént bemutatott statisztikai mutatókat az adathalmaz PostgreSQL-ben tárolt részéről (míg a **3. táblázat** tartalmazza a MySQL adatbázis kimutatásait). Kiemelném még a befejezett játékok átlagos kattintásszámait is, melyek mindkét esetben 6 körüli értéket vesznek fel – azaz a Wikipédia is hasonló tulajdonságú kisvilágnak tűnik, mint amilyenek Milgram kísérletében a társadalom bizonyult (Milgram, 1967).

2. táblázat A The Wiki Game adatbázis PostgreSQL alapú részének tulajdonságai

PostgreSQL	
Tulajdonság	Érték
Játékban megjelent cikkek száma	1.164.592
Játékok száma (ebből befejezett játékok)	9.179.157 (3.259.019)
Átlagos játékidő (befejezett játékok esetén)	65,5 s
Maximális játékidő	93,11 s
Kattintások átlagos száma az összes (és a befejezett) játékokra	5,663 (6,53)
Minimum és maximum kattintás	1 és 254
Játékosok száma	649.737
Legalább 30 (50, 100, 500) befejezett játékot lejátszó játékosok száma	17.573 (7.097, 1.973, 121)

3. táblázat: A The Wiki Game adatbázis MySQL alapú részének tulajdonságai

MySQL	
Tulajdonság	Érték
Játékok száma (ebből befejezett játékok)	5.416.783 (2.100.055)
Átlagos játékidő az összes (és a befejezett) játékok esetén	525,72 s (98,36 s)
Kattintások átlagos száma	6,98
Minimum és maximum kattintás	1 és 2.777
Játékosok száma	128.487
Legalább 30 (50, 100, 500) befejezett játékot lejátszó játékosok száma	16.230 (8.429, 2.878, 141)

4.2.2 A Wikipédia hálózat tulajdonságai

Maga a Wikipédia, a játék mögött álló hálózat folyamatosan fejlődő és növekvő gráf, angol kiadása⁸ az idei év őszén 6,2 millió szócikket számlált. Gyors változása miatt nem érhetőek el olyan friss elemzések, amelyek hálózati tulajdonságairól naprakészen számolnának be, viszont mivel fejlődése organikus és a kezdetek óta változatlan szabályszerűségek mentén zajlik, feltehetjük, hogy jellege nem változott meg a régebbi elemzések elvégzése óta, így azokat fenntartások nélkül vehetjük alapul, midőn a hálózat elemzésünkhöz szükséges, fontos tulajdonságait sorra vesszük. A Wikipédia rendelkezik mindazokkal a tulajdonságokkal, amelyeket az irodalmi áttekintésben a komplex hálózatok általános vonásaiként bemutattam (Helic, 2012). Bár mérete hatalmas, ahhoz képest kicsi, 6 kattintás körüli átmérővel rendelkezik, következésképpen elmondható, hogy rendelkezik a kisvilág tulajdonsággal. Irányított gráf, így a megfelelő fokszámeloszlásokat szét kell bontani a ki- és be fokszámoknak megfelelően, s az így felállított eloszlások középső részükön hatványfüggvény eloszlást mutatnak $\alpha = 2 \dots 3$ -mas értékkel, e nagyfokú heterogenitással teret nyitva a szupercsúcsok és a hosszú nyúlványok létrejöttének. Klaszterezettségét tekintve a gráf lokálisan erősen kapcsolt, C értéke 0.3 körül van⁹. A tulajdonságok hasonló kombinációja alapján elmondható, hogy a Wikipédia előnyös környezetet teremt az emberi navigációhoz (Boguna, 2013).

4.2.3 Az útvonalak nyúlásának vizsgálata

Egy-egy játékos útvonalait az adatbázisok *game_game* és *game_click* nevű táblái tartalmazták. Annak érdekében, hogy a feldolgozandó adatmennyiség kezelhető legyen, első lépésben egy SQL lekéréssel kiszűrtem az adathalmazból a számomra fontos mezőket (a játék kezdő és cél cikke, a játékos azonosítója, a kattintások száma és a játékidő), majd az adatokat játékosonként külön elemeztem. Hogy az általuk megtett utakat összehasonlíthassam az optimális (legrövidebb) útvonalakkal, fel kellett építenem a teljes angol Wikipédia linkhálózatát, ami nem bizonyult egyszerű feladatnak.

⁸ A The Wiki Game és így az én vizsgálatom is kizárólag az angol nyelvű Wikipédiát veszi figyelembe.

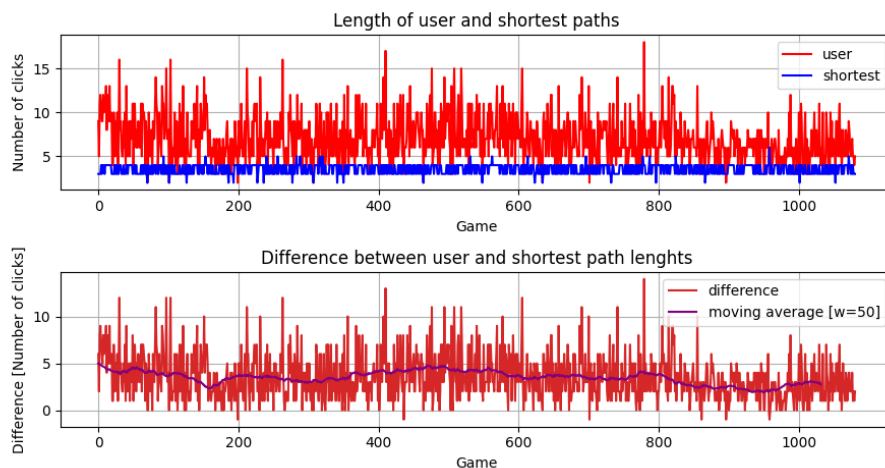
⁹ A konkrét szám adatok függenek a Wikipédia gráfrepresentációjának eltérő megközelítéseitől, amelyeket a kutatócsoportok alkalmaztak.

Első megközelítésben egy letisztított, funkcionális adatbázis (ami kizárólag a Wikipédia lapjait és a köztük mutató linkeket tartalmazza (Consonni, 2019)) alapján az *igraph* programcsomag segítségével kezdtem bele a gráf rekonstrukciójába, ám ez a megoldás rövid idő alatt használhatatlannak bizonyult, ugyanis a több száz millió linket tartalmazó adatszett feldolgozása egy általános számítógépen több mint hatvan napig tartott volna.

További internetes kutatás után ráakadtam olyan szoftvermegoldásokra, amelyek képesek a Wikipédia két cikke között legrövidebb utakat keresni (többek között a *Six Degrees of Wikipedia*, a *WikiRacer* és a *degreesofwikipedia.com* is ilyen programok), és ebben a feladatban értékelhetően gyorsnak is bizonyulnak. Mivel nem bocsátottak ki bárki számára hozzáférhető online API-t¹⁰, végül az elsőként említett szoftver lokálisan futó példánya mellett döntöttem, ami egy lekérést 0,5-1 másodperc alatt képes már teljesíteni.

Python nyelven írt elemző szkriptem a felhasználó minden egyes játékára kiszámítja a legrövidebb utat (ahol több ilyen is van, azok közül véletlenszerűen választ egyet¹¹), majd az eredményekből statisztikát készít. Hogy képet kaphassunk a játékos saját maga és a legrövidebb útkereső algoritmus által használt élekről és az ezekben esetleg megfigyelhető szabályosságokról, a legrövidebb utakból élhasználati gráfot készítettem, amelyet a következő szakaszban a törzshálózat ellenpontjaként mutatok be.

6. ábra: Egy játékos saját útjai és az optimális utak (fent), illetve ezek differenciája



Forrás: saját munka

A **6. ábra** egy játékos eredményeit mutatja be az útvonalak hosszát illetően. Ő az útvonalakat 7,04 lépéses átlagos hosszal és 2,57 szórással teljesítette. Látható, hogy teljesítménye az esetek túlnyomó többségében nem éri el az átlagosan 3,57 lépésben sikeres BFS (Cormen, 2001) algoritmusét, hanem középértékben 3,47 lépéssel többbe kerül eljutnia a célhoz. Ez a hátrány az összes elemzett személy között tipikusnak mondható, bár előfordultak olyan emberek, akik a hátrányukat 3 lépés körüli differenciáig dolgozták le.

¹⁰ alkalmazásprogramozási interfész, melynek segítségével egyszerű HTTP kéréseken keresztül oldhatóak meg a felhőben számításigényes feladatok

¹¹ Mivel az algoritmus BFS keresést használ, az útválasztásban megjelenő determinizmus manipulálhatná az élhasználati statisztikákat.

Hasonlóan a *fit-fat-cat* kutatáshoz (Gulyás, 2020), a karakterisztikus nyúlás a teljes mintán megjelenik és mértéke kis ingadozásokkal minden személyre hasonlóan tűnik. Nem, illetve rendkívül kis mértékben jelenik meg azonban a 'tanulási periódus' vége után bekövetkező javulás az úthosszak tekintetében.

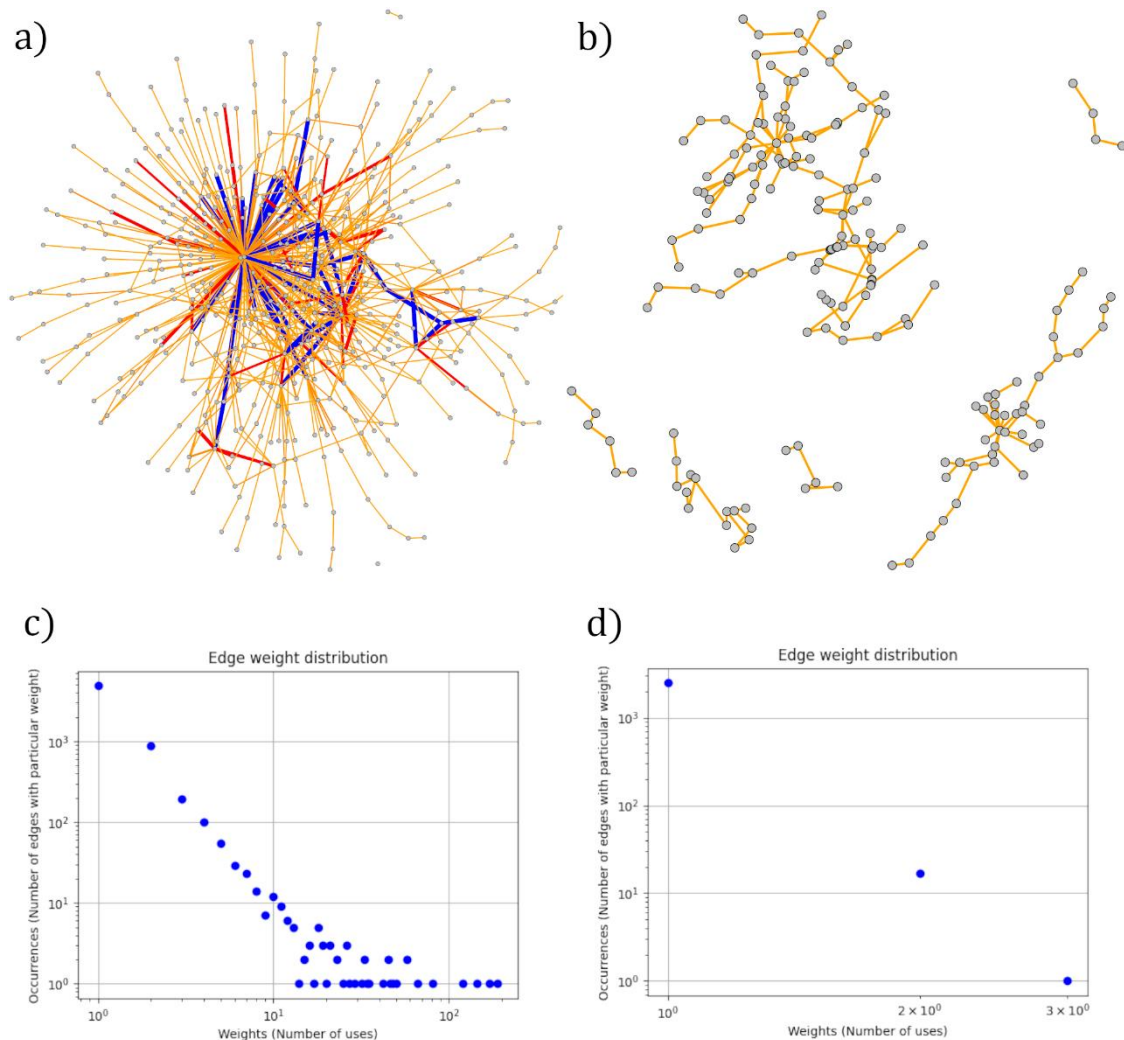
Az idézett kutatásban a játék hálózatán navigáló egyének első játszmáik alatt míg a hálózatot kiismerték (valamint a törzshálózatot felépítették), rendre rosszabb eredményeket értek el, míg egyre több menet lejátszása után rohamosan fejlődtek. Meglátásom szerint ez a jelenség a Wikipédia esetében két okból nem figyelhető meg: egyrészt a navigációhoz szükséges képességbeli feltételek már az első játszma alkalmával kielégítettek (mivel a játékhoz csak a 'kattintás' gyakorlottsága és a linkek felismerésének képessége szükségeltetik, amit minden számítógépfelhasználó készségszinten ismer); másrészt a Wikipédia szemantikus tartalma által alkotott rendszer nagyon hasonlóan épül fel az ember kognitív tudásához képest (a Wikipédián a macska oldal feltünteti, hogy a faj a macskafélék családjába tartozik, ugyanez a tudás az emberi elmében is kialakult már gyermekkorában).

4.2.4 Törzshálózat kialakulásának vizsgálata

Hasonlóan az útvonalak nyúlásának vizsgálatához, ebben az esetben is szükség volt az adatok előszűrésére: miután a törzshálózat a játékos által leggyakrabban választott élekkel áll összhangban, minden egyes kattintást bele kellett foglalnom az élhasználati gráf felépítésébe. Ezért a *game_click* táblák adatait vettem alapul, kiszűrve közülük a számomra haszontalanokat, és megtartva a kattintás célját, idejét, a kattintó játékos azonosítóját és annak a játéknak az azonosítóját, melyhez a kattintás tartozik. A kattintásokból rekonstruáltam a felhasználók élhasználati gráfjait, kiszámítottam e gráfok fontos paramétereit (mint például az átmérő és átlagos úthossz) és végül ábrázoltam őket. Az ábrázolás során a gráfok nagy mérete miatt (a legtöbb játékos esetén elérték a több ezer csúcsot), és hogy az eredmények szembeütőbbek legyenek elhagytam a közbeeső centralitás szerint az alsó 90 percentilisbe tartozó csúcsokat, így kvázi a hálózat 'belső', fontos része látszik az ábrán.

Szemléltetésként egyetlen példajátékos eredményei láthatóak a **7. ábrán**, ám ezek minden játékos esetén hasonlóak. Jól megfigyelhető a kialakuló törzshálózat az **a)** részben, ahol az egyre sötétedő színek egyre fontosabb éleket jelölnek. Az ábra bal alsó felén az élek használat szerinti eloszlása látható, mely roppant érdekes módon hatványfüggvény eloszlást követ, szélein a várható eltérésekkel, tehát vannak olyan élek, melyeket csupán egyszer-kétszer járt végig a felhasználó, míg előfordulnak olyanok is, melyek rengeteg út részeként szerepeltek, azokat gyakran választotta a navigáció során.

7. ábra: Egy játékos törzshálózatának vizsgálata: a) az élhasználati gráfon kirajzolódó törzshálózat b) ugyanezekre az utakra a legrövidebb utat megtaláló algoritmus élhasználati gráfja c) a játékos élhasználatának eloszlása és d) az algoritmus élhasználatának eloszlása kétszeres logaritmikus skálán



Forrás: saját munka

Kirajzolva az élhasználati gráf fokszámoszlását is kétszeres logaritmikus skálán, hasonló eloszlást tapasztalunk: azaz a gráf skálafüggetlen tulajdonságúnak tűnik, és nagyfokú heterogenitás jellemzi. Meg is jelennek a fokszámoszlás által jóslott hosszú láncok, és a kis számú, de annál számosabb kapcsolattal rendelkező *hubok* (szupercsúcsok) is.

Mindent egybevéve megállapítható, hogy a játékos a navigáció során nem egyformán használja a rendelkezésére álló lehetőségeket, hanem kialakít egy gyakrabban használt, bejáratos hálózatrészt, amely segítségére van a feladat végrehajtásában: ezt nevezzük törzshálózatnak. Összehasonlításként az ábra jobb oldalán lévő **b)** rész tartalmazza ugyanezen start-cél cikkipárok közötti legrövidebb utakból összeállított élhasználati gráf közbeeső centralitás szerinti felső 10 percentilisbe tartozó részét. Világosan megfigyelhető, hogy ebben a gráfban az élek kihasználtsága jóval alacsonyabb, azok eloszlása közel homogén és maga a centrum is több komponensre esik szét. Egy pillantást vetve az élhasználati eloszlására is, ami a

d) részben figyelhető meg, látszik, hogy nem fordulnak elő kitüntetett használatú élek és utak. Elmondható tehát, hogy sokkal kevésbé centralizált és nem sokban különbözik a véletlen gráfoktól.

Egyes játékosok vizsgálata során az adathalmaz túlnyomó részében rögzíthettem, hogy kialakult a játékosok törzshálózata – avagy annak kezdeménye –, ami fontos eredmény tekintve a Wikipédia hatalmas méretét és ehhez mérten a játszmák számának elenyésző voltát. Úgy tűnik tehát, hogy a hipotézisben felvetett állítás megerősödik: az egyes játékosok nem egyforma eséllyel válogatnak a Wikipédia oldalak között egy-egy útkeresés során. Vannak jeles pontok, melyek a teljes hálózat egy, a játékos elméjében létrejövő (vagy már addig is létező) egyszerűsített hálózatnak a részei, s a navigáció során a részhálózatban lévő útvonalak előnyben részesítettek a többi útvonalhoz képest. Ezáltal a navigáció teljes folyamata leírható a *'juss el a törzshálózattig, érij a cél közelébe, menj a törzshálózattól a célig'* minta alapján. Valószínűsíthető emellett, hogy a navigációt hierarchikus szabályok is meghatározzák, a több tanulmányban is leírt *'tarts a hálózat közepe felé, majd ereszkedj a célig'* séma alapján (Leskovec, 2012; Helic, 2012; West, 2009), erre utalnak az élhasználati gráfban megjelenő centrikus szupercsúcsok is, melyeken keresztül nagyságrendekkel több áthaladás történik, mint egyéb csúcsokon.

4.2.5 A Matrice adatbázis elemzése - az adatbázis és az adatok szerkezete

A Matrice Android alkalmazás által gyűjtött adatok a Firebase alkalmazásfejlesztési platform által összegyűjtve és egységes formába öntve álltak rendelkezésemre. Nagy előnye a használt JSON adatstruktúrának, hogy mind emberi, mind "gépi" szemmel egyszerűen olvasható. Az egy játékosról tárolt adatok automatikusan egy anonimizált felhasználói kulcshoz lettek rendelve, s tartalmazzák a lejátszott játékok adatait¹² és olyan, a játékosra vonatkozó statisztikai információkat, mint például annak neve és kora. Egy játékról ezen belül a következő rekordokat tartjuk számon: a játéktábla méretét (minden játszma a 3x3-as táblán zajlott), a kezdő és végső állapotok kódját, a megtett útvonal hosszát, a játékidőt, egy időbélyeget és a megtett út során érintett állapotkódok felsorolását.

A Matrice adatbázis 74 játékosról tartalmaz adatokat¹³. Előfeldolgozásként kiszűrtem ezek közül azokat, akik 35 játszmánál kevesebbet játszottak le, mert róluk így nem áll rendelkezésre elég adat a következtetések helyes levonásához, ezáltal 32 játékos maradt. Feltételeket állítottam az egyes játszmákra vonatkozóan is: elvettem az adatok közül azokat, amelyek több, mint harminc lépésből álltak. Feltételeztem, hogy ezekben az esetekben a felhasználó nem tudatos navigációt végzett, hanem céltalanul bolyongott, mivel ekkora lépésszám a gráf átmérőjét hatszorosan meghaladja. Ugyanezen okból elvettem azokat is, ahol a játékidő szökött 100 másodperc fölé, másik valószínű magyarázat, hogy itt a felhasználó félbehagyta a játékot és csak később tért vissza hozzá. Mindezen szűréseket elvégezve 3.697 játszma maradt benne az elemzési halmazban.

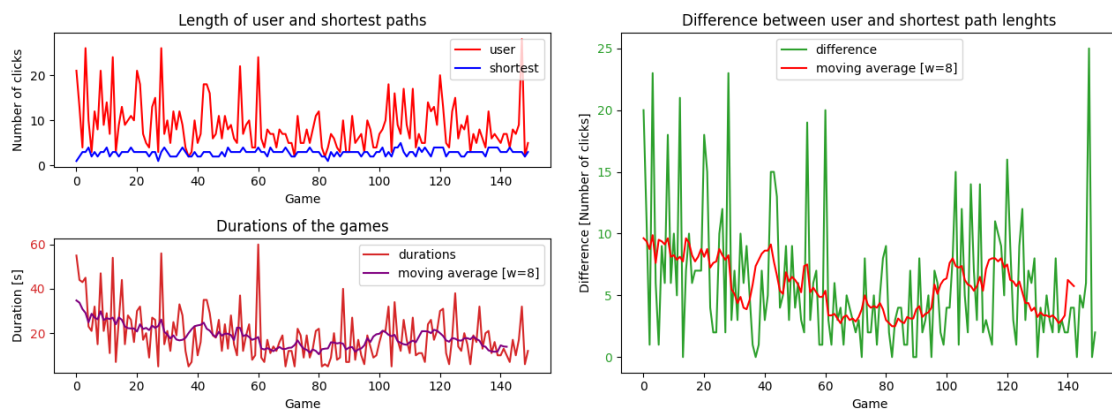
¹² egyedül a befejezett játékokét

¹³ Mint a 4. bekezdésben leírtam, a regisztrált felhasználók száma 85 volt, közülük 11-en valószínűleg internetkapcsolat nélkül vagy egyáltalán nem játszottak a játékkal.

4.2.6 Az útvonalak nyúlásának vizsgálata

A játékosok útvonalainak vizsgálata nagyon hasonló módon zajlott a **4.1.4** pontban ismertetett módszerhez. Könnyebbséget jelentett, hogy a Matrice adathalmazban minden játszmahoz már egyben rendelkezésre állt az útvonal. Mivel itt a teljes gráf sem volt kezelhetetlen méretű (lásd az **1. táblázat** adatait) a legrövidebb utak megkereséséhez saját programkódot írtam. Szintén a kódban összesítettem és jelenítettem meg a legrövidebb utak és a játékos által bejárt valódi utak statisztikáit. Hasonlóan az előzőekhez minden játékosról egyéni analízis készült (bár szerencsére ennél az adatbázisnál az adatok limitált sokasága következtében a programkódok rövid idő alatt lefutottak).

8. ábra: Egy játékos saját útjai és a megfelelő optimális utak (balra fenn), a játékidő alakulása (balra lenn) és a differencia a saját és optimális utak között (jobbra)



Forrás: saját munka

A **9. ábra** jeleníti meg egy kiválasztott játékos útadatait. Ez a felhasználó átlagosan 8,85 lépés alatt oldotta meg a feladványokat (5,39 lépés szórással), hozzá képest a Dijkstra (Cormen, 2001) algoritmus 2,9 lépést tett a starttól a célig, azaz középértéken 5,96 lépéssel megelőzte emberi vetélytársát. Kétségtelenül nagy a nyúlás mértéke, körülbelül megegyezik a sokkal nagyobb Wikipédián megfigyelt nyúlással, enyhén még meg is előzi azt. Ugyanakkor itt is kijelenthető, hogy a nyúlás hasonló értékkel az összes vizsgált személynél megfigyelhető.

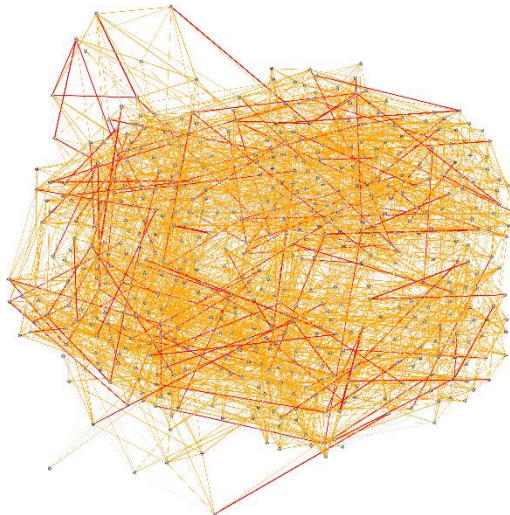
Ha szemügyre vesszük a játékidőket tartalmazó grafikon, szembe tűnhet (alaposabb tanulmányozás után még a differencia ábrázoláson is) a Wiki Game adatbázis elemzése esetében elmaradt tanulási jelenség. A kezdetben rendre 40 másodperc körüli eredményeket elérő játékos időátlagát az adatsor vége felé egészen 20 másodperc közelébe szorította. Ugyanez a hatás vehető észre a megtett és a legrövidebb utak különbségein is, bár sokkal kisebb mértékben. Összhangban az ott kifejtett érvekkel, a jelenség feltűnését a Matrice játék nem triviális irányíthatóságának tulajdonítom¹⁴.

Még egy érdekes hatást szeretnék megmagyarázni: a grafikonon olykor-olykor megjelenő idő- és lépésszámbeli kiugrásokat, tüskéket. Feltevésem szerint ezekben az esetekben a játékos gyökeresen más, bolyongó „stratégiát” használt, ami nem nevezhető teljes joggal stratégiának, mivel nem áll másból, minthogy a logikai fonalat elvesztett személy

¹⁴ Természetesen az a kritérium is teljesül, hogy az ember ismeretlen hálózattal találja szembe magát, viszont a következő szakaszban ismertetett eredmények fényében nem bizonyos, hogy ebben a hálózatban a nyúlás változása szempontjából ez a faktor releváns.

véletlenszerűen teszi meg lépések egymásutánját, mígnem ismert helyre (vagy annak közelébe lyukad ki); ennek a stratégiának létezését már mások is felvetették (Riascos & Mateos, 2012).

9. ábra: Egy játékos élhasználati gráfja (színezett élek) a Matrice gráfra vetítve (szürke élek



Forrás: saját munka

4.2.7 A törzshálózat kialakulásának vizsgálata

Játékosok által kialakított törzshálózat vizsgálatánál a **4.1.5** szakaszban bemutatott módszert követtem. Itt a hálózat mérete lehetővé tette, hogy a szemléltetés során a használt éleket az eredeti gráfra vetítve jelenítsem meg. Így a felhasználói útvonalakat lépésenként feldolgozva minden lépésben egyel megnövelt súlyt rendeltem az éppen bejárt élhez. Ebben a vizsgálatban figyelembe vettem az előző részben elhanyagolt hosszú játszmák éleit is, hiszen itt egyedül annak van jelentősége, hogy az ember azt meghatározott élt választja egy döntési helyzetben; annak, hogy ez milyen játszmában történik, immár nincs.

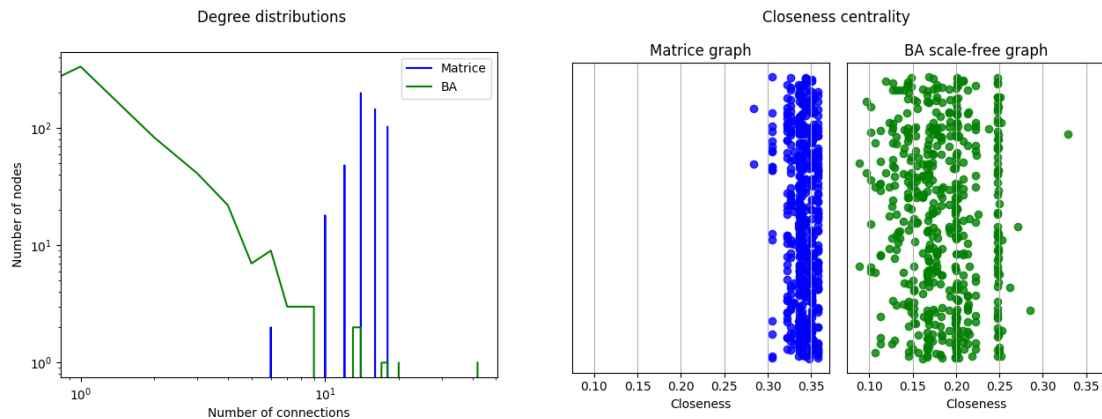
Az eredményeket összesítve kijelenthető, hogy a Matrice játék esetén nem alakul ki a törzshálózat. Megfigyelhető a **9. ábrán** (és az összes többi személy élhasználati gráfján), hogy a játékos látszólag egyenletesen és hierarchikus szabályszerűségek követése nélkül használja fel a gráf egyes éleit. Felmerül tehát a kérdés, hogy szemben a The Wiki Game esetével, mi okozza a törzshálózat kialakulásának elmaradását?

Többféle lehetséges magyarázat is kínálkozik, melyek rávilágítanak a két hálózat különbözőségének egyes aspektusaira. Megközelíthetjük a kérdést a rendszerek tartalmának oldaláról: mind a Wikipédia, mind pedig a referencia kutatásban vizsgált fit-fat-cat játék szóhálózatának csomópontjai szemantikus jelentéstartalmakat hordoznak, emellett pedig olyan komplex rendszerek részei, mint az általános kognitív sémarendszer, illetve maga a nyelv, ezáltal az emberek számára már ismerősek, belakottak és gyakorlottak. Ezzel szemben a Matrice játék hálózatának pontjai absztrakt mintákat jelenítenek meg, amelyek között elvont (bárha egyszerű) logikai szabályrendszer teremt kapcsolatot, ráadásul az emberi elme számára bár felfogható, semmiképpen sem ösztönös. Ez a különbség okozója lehet annak, hogy a bevált navigációs stratégia helyett az emberi elme valami újat keres.

Második (és párhuzamos) magyarázatként érdemes megvizsgálni a Matrice mögött álló hálózat tulajdonságait alaposabban. Mivel a gráf teljes mértékben mesterséges szabályok alapján alakult ki (ellentétben a szó- vagy a fogalmi hálózatokkal), nem önszerveződő

formában, ezért várhatóan tulajdonságai merőben eltérnek a természetben előforduló komplex hálózatok tulajdonságaitól. Valóban, a Matrice gráf kis lokális klaszterezettséggel rendelkezik, emellett a **10. ábrán** látható, hogy fokszámeloszlása sokkal inkább a véletlen gráfok Poisson-eloszlásához hasonló, mint a skálafüggetlenek hatványfüggvény eloszlásához.

10. ábra: A Matrice hálózat és egy azonos méretű Barabási-Albert gráf összehasonlítása fokszámeloszlás (balra) közelségi centralitás szerint (jobbra)



Forrás: saját munka

Hogy az eltérés még nyilvánvalóbbá váljon, a Matrice gráf mellett ábrázoltam egy méretben megfelelő, ám Barabási-Albert modell szerint szerveződő gráf tulajdonságait. Az ábra jobb oldalán megjelenítettem a két hálózat csúcsait közelségi centralitás szerint is.

Látható, hogy a Matrice hálózat csúcsai körülbelül azonos értéket vesznek fel, és az értékek szórása roppant kicsi – mintha minden csúcs a hálózat középpontjában lenne; míg a BA modell egyenlő arányban tartalmaz periférikus és félperiférikus csúcsokat, s igen kis részük található csupán a hálózat magjában. Ez szemléletesen azt jelenti, hogy a Matrice hálózatnak nincs igazi közepe, nincs valódi perifériája, csúcsai nagyon nagy mértékben hasonlóak egymáshoz. Nem csoda, hogy a tájékozódáshoz nem nagyon találunk benne fogódzót, hiszen ez a hálózat nem segíti azt. Ellenkezőleg: a mindenfelé tapasztalható homogenitás miatt elveszítjük helyzetérzékelésünket. Ahelyett, hogy törzshálózatot próbálnánk felépíteni, elvont logikai szabályokra kell támaszkodnunk, más stratégiát kell alkalmaznunk a navigációnk során.

Játékosok személyes megkérdezése után egy stratégiát vázoltok fel (hogyan ezt az adatokból milyen módon lehetne kimutatni, egy jövőbeli tanulmány kérdése): többen követték a szabályt, hogy elsőként invertálás lépésekkel elérték a nullák és egyesek célnak megfelelő számát a táblán, majd forgatás művelettel azokat a helyükre tolták. Hogy ez a taktika mennyire jó, azt nehéz megállapítani, annyi azonban biztosan kijelenthető: a legrövidebb utakat csak szerencsés esetekben találja meg.

Összefoglalva, bár a Matrice gráfban is megjelenik az útvonalak megnyúlása, törzshálózat nem alakul ki benne. Tulajdonságai alapján nem is alkalmas a navigáció ezen formájának támogatására. Az a tény, hogy az emberek itt is képesek valamilyen szuboptimális tájékozódásra más stratégiák eredménye.

5. Összefoglalás

Kutatásom során az emberi navigációt próbáltam jobban megérteni komplex hálózatokban. Ehhez olyan játékokat hívtam segítségül, amelyek mögöttes logikai szerkezete egy komplex hálózatra, játékmódja pedig a hálózatban végrehajtandó navigációs feladatra fordul le. Hipotézisem szerint az emberek által használt útvonalak karakterisztikusan megnyúlnak, mivel az emberek óhatatlanul igyekeznek optimalizálni az elméjüket terhelő információmennyiséget szemben az útvonalak optimalitásával. Két játékból, az önszerveződő, természetes tulajdonságokat mutató Wikipédia cikkek kapcsolati hálózatából és a mesterséges, véletlen gráfokhoz közeli Matrice gráfból nyert adatok segítségével vizsgáltam meg a hipotézis érvényesülését. Eredményeim igazolták, hogy a Wikipédia előnyös környezetet biztosít a benne való tájékozódáshoz, az emberek egyszerűsített törzshálózatok mentén folytatják navigációjukat. Ezzel ellentétben a Matrice hálózat nem teremt ideális körülményeket, nagymértékben uniform, ami a helyzetérzékelést lehetetlenné teszi. Nem is alakul ki benne törzshálózat, a nyúlásokból ugyanakkor más stratégiák jelenlétére következtethetünk.

Érdekes lehet a jövőben megvizsgálni a Wikipédiát leegyszerűsítő törzshálózatok által elért megtakarítást az információ tárolásában, ehhez a fit-fat-cat kutatás szolgáltathat követendő módszert (Gulyás, 2020), illetve a nyúlás függését ennek méretétől. Szintén érdekes, hogy a Matrice hálózat esetében vajon milyen stratégiát követ az emberi elme, és hogy ebben az esetben milyen tényezők karakterizálják az útvonalak megnyúlását.

Irodalomjegyzék

- Albert, R., Barabasi, A-L. (2001). *Statistical Mechanics Of Complex Networks*. Reviews of Modern Physics. 74. 10.1103/RevModPhys.74.47.
- Barabási A-L., Pósfai, M. (2016). *Network Science*. Cambridge: Cambridge University Press
- Black, P.E. "greedy algorithm", in *Dictionary of Algorithms and Data Structures [online]*, ed. 2 February 2005. Letöltés dátuma: 2022.04.27. Forrás: <https://www.nist.gov/dads/HTML/greedyalgo.html>
- Boguña, M. & Krioukov, D. & Claffy, K. (2007). *Navigability of Complex Networks*. Nature Phys.. 5. 10.1038/NPHYS1130.
- Clemesha, A.. *The Wiki Game*. Letöltés dátuma: 2022.04.27. Forrás: <https://www.thewikigame.com>
- Consonni, C., Laniado, D. & Montresor, A. (2019). *WikiLinkGraphs: A complete, longitudinal and multi-language dataset of the Wikipedia link networks*.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2001). *Introduction to algorithms*. (Second ed.). MIT Press and McGraw–Hill. pp. 595–601. ISBN 0-262-03293-7.í
- Csardi G., Nepusz T. (2006). *The igraph software package for complex network research*. InterJournal, Complex Systems, 1695. <https://igraph.org>
- Csertán T. (2020a) *Hálózati játék adatok, elemzőkódok a Githubon*. Letöltés dátuma: 2022.04.27. Forrás: https://github.com/csertant/network_games_analysis
- Csertán, T. (2020b). *Matrice kód a Githubon*. Letöltés dátuma: 2022.04.27. Forrás: <https://github.com/csertant/matrice>
- Csertán, T. (2020c). *Matrice*. Letöltés dátuma: 2022.04.27. Forrás: <https://play.google.com/store/apps/details?id=com.nosebite.matrice>
- Csoma, A., Kőrösi, A., Rétvári, G. et al. (2017). *Routes Obey Hierarchy in Complex Networks*. *Sci Rep* 7, 7243. <https://doi.org/10.1038/s41598-017-07412-4>
- Gulyás, A., Bíró, J., Rétvári, G. et al. (2020). *The role of detours in individual human navigation patterns of complex networks*. *Sci Rep* 10, 1098 <https://doi.org/10.1038/s41598-020-57856-4>
- Helic, D. (2012). *Analyzing user click paths in a Wikipedia navigation game*. MIPRO, 2012 Proceedings of the 35th International Convention. 374-379.
- Kleinberg, J. (2000). *Navigation in a small world*. *Nature*. 406. 845. 10.1038/35022643.
- Leskovec, J. (2012). *Human navigation in networks*. HT'12 - Proceedings of 23rd ACM Conference on Hypertext and Social Media. 10.1145/2309996.2310020.
- Milgram, S. (1967). *The small world problem*. *Psychology Today* 2, 60–67
- Newman, M. (2010). *Networks: an introduction*. Oxford: Oxford university press.

- Newman, M.E.J. (2003). *The Structure and Function of Complex Networks*. Computer Physics Communications. 147. 40-45. 10.1016/S0010-4655(02)00201-1.
- Pérez Riascos, A. & Mateos, J. (2012). *Long-range navigation on complex networks using Lévy random walks*. Physical review. E, Statistical, nonlinear, and soft matter physics. 86. 056110. 10.1103/PhysRevE.86.056110.
- Watts, D., Dodds, P. & Newman, M.E.J. (2002). *Identity and Search in Social Networks*. Science. 296. 1302-. 10.1126/science.1070120.
- West, R., Pineau, J. & Precup, D. (2009). *Wikispeedia: An Online Game for Inferring Semantic Distances between Concepts*. IJCAI International Joint Conference on Artificial Intelligence. 1598-1603.
- Zhu, S., Levinson, D. (2015). *Do People Use the Shortest Path? An Empirical Test of Wardrop's First Principle*. PLOS ONE 10(8): e0134322. <https://doi.org/10.1371/journal.pone.0134322>