

Kovács Bence

A 2024. ÉVI LABDARÚGÓ EURÓPA- BAJNOKSÁG KIÉRTÉKELÉSE PÁROS ÖSSZEHASONLÍTÁSI MÓDSZEREKKEL

Témavezető: Dr. Mihálykóné dr. Orbán Éva

Műszaki Informatikai Kar, Matematika Tanszék

Pannon Egyetem
Intézményi Tudományos Diákköri Konferencia

2024.



A 2024. évi labdarúgó Európa-bajnokság kiértékelése páros összehasonlítási módszerekkel

*Készítette: Kovács Bence
Programtervező informatikus BSc
Műszaki Informatikai Kar, Matematika Tanszék*

Témavezető: Dr. Mihálykóné dr. Orbán Éva

A legtöbb sporteseménynél, mint ahogy a labdarúgásnál is, a rangsorok elsődlegesen kumulatív pontszám alapján kerülnek kialakításra, ahol pontokat egy csapat a lejátszott mérkőzések során kaphat, amennyiben döntetlent játszik, vagy legyőzi ellenfeleit. Ennek a kiértékelési módszernek az egyik legnagyobb hátránya, hogy nem veszi figyelembe az egymás ellen játszó csapatok erősségét.

A páros összehasonlítási módszerek azonban képesek ennek figyelembe vételére, továbbá képesek még különböző mértékű győzelmeket is feldolgozni. A kiértékelések eredményeként olyan csapatok között is kaphatunk sorrendet, amelyek nem játszottak egymás ellen. Ezeknél a módszereknél viszont a kiértékelhetőségnek több előfeltétele van.

A dolgozatban a 3 és 5 kategóriás Bradley-Terry modell használatával kerülnek egymáshoz viszonyított kiértékelésre a 2024. évi UEFA labdarúgó Európa-bajnokság nemzeti csapatai. Az irodalomban használt 3 kategóriás modellről vizsgálataim során megállapítottam, hogy nem alkalmas az összes lejátszott mérkőzés alapján a sorrend és az erősségek meghatározására, így áttértem a több információt feldolgozó 5 kategóriás modellre. Ebben azonban meg kell állapodni abban, hogy mely gólkülönbség alapján soroljuk a mérkőzést a nagy győzelmek, nagy vereségek illetve a kis győzelmek, kis vereségek közé. A paraméterbecslésre maximum likelihood becslést alkalmaztam. A végrehajtására parciális deriváltak zérushelyeit megkereső fixpont iterációs numerikus optimalizálási módszert dolgoztam ki az 5 kategóriás esetben. A kapott kiértékelési eredményeket összehasonlítottam egy, a továbbjutások és a csoportkörökben megszerzett pontok alapján számított sorrenddel.

A kidolgozott módszer és az elkészített számítógépes programok jól használhatók más esetekben is, akár sportbajnokságok eredményeinek kiértékelésére, akár egyéb más páros összehasonlításból származó adatok elemzésére.

Kulcsszavak: páros összehasonlítási modellek, Bradley-Terry módszer, numerikus eljárások, sportmérkőzések kiértékelése

Köszönetnyilvánítás

Köszönetet szeretnék mondani Dr. Mihálykóné dr. Orbán Évának, mivel témavezetőként elősegítette a pályamunkám létrejöttét, továbbá Dr. Mihálykó Lajos Csabának, mivel témavezetői kötelezettség hiányában is értékes visszajelzést adott a munkám minőségéről.

Tartalomjegyzék

1. Bevezetés	6
2. Sztochasztikus háttérű páros összehasonlítási modellek	10
2.1. Az általánosított Bradley-Terry modell	10
2.1.1. 3 kategóriás modell	11
2.1.2. 5 kategóriás modell	16
3. Kiértékelhetőségi feltételrendszerek	23
3.1. Feltételrendszer 3 kategóriás modell esetén	23
3.2. Feltételrendszer 5 kategóriás modell esetén	26
4. A 2024. évi UEFA labdarúgó Európa-bajnokság eredményei- nek kiértékelése	28
4.1. A mérkőzések eredményei	28
4.2. Hagyományos kiértékelések eredményei	31
4.3. A mérkőzések eredményeinek kiértékelése 3 kategóriás modellben.	34
4.3.1. A mérkőzések eredményeinek kiértékelhetősége	34
4.3.2. A kiértékelhető részhalmaz eredményei a 3 kategóriás modellben	43
4.3.3. Egyes kiértékelési eredmények közötti korrelációk	48
4.4. A mérkőzések eredményeinek kiértékelése 5 kategóriás modell alkalmazása esetén	51
4.4.1. Eredmények különböző súlyozások alkalmazása esetén	51
4.4.2. A különböző kiértékelési módszerek eredményeinek korrelációja	66
5. Informatikai megvalósítás	70

6. Összefoglalás **76**

Irodalomjegyzék **78**

1. fejezet

Bevezetés

Páros összehasonlítási módszereket gyakran alkalmaznak különböző területeken, például döntéshozatalban [1], oktatásban [2], informatikában [3, 4, 5], de páros összehasonlítási eredményekként interpretálhatjuk sportmérkőzések eredményeit is [6]. Használatuk kimondottan előnyös olyan esetekben, amikor a skálák értelmezése nagyon szubjektív. Nehéz ugyanis számmal jellemezni az objektumokat bizonyos szempontok szerint (például fontosság, szépség, súlyosság, stb.) de biztosabban lehet véleményt mondani az objektumokról, ha egymással hasonlítjuk össze őket, vagyis arról alkotunk véleményt, hogy ketjük közül melyik fontosabb, szebb, súlyosabb stb.

Fejlődésük is pszichológiai területről indult körülbelül 100 éve. Thurstone az 1920-as években publikálta azt az ötletét, hogy az egyes összehasonlítandó objektumok mögött véletlen mennyiségek, valószínűségi változók húzódnak meg, és az objektumok összehasonlításának eredménye a látens valószínűségi változók különbségétől függ [7]. Ezeket a páros összehasonlítási módszereket sztochasztikus háttérű páros összehasonlítási módszereknek nevezzük. Thurstone 2 kategóriát (opciót) engedett meg a döntések során, jobb és rosszabb kategóriákat, valamint normális eloszlású valószínűségi változókkal dolgozott. Bradley és Terry normális eloszlás helyett logisztikus eloszlást feltételezett a különbségekre, ugyancsak két kategóriát megengedve a döntések során [8]. Thurstone a paraméterek becslésére a legkisebb négyzetek módszerét alkalmazta, míg a Bradley-Terry modellben maximum likelihood becslést használtak. A kétkategóriás modelleket később általánosították három kategóriára, bevezetve az egyforma opciót [9, 10]. Davidson tovább módosította ezen modelleket mértani közép tulajdonságot feltételezve az egyforma és a jobb/rosszabb döntés

valószínűségei között [11]. Három kategóriánál több opciót megengedő modell található a [12] és a [13] publikációkban legkisebb négyzetes illetve maximum likelihood paraméterbecslési módszerekkel.

Kétkategóriás Bradley-Terry modell esetén Ford adott szükséges és elégséges feltételt a maximum likelihood becslés létezésére és egyértelműségére, valamint fixpont-iteráción alapuló numerikus módszert mutatott a maximum helyének megkeresésére [14]. Hasonló numerikus módszerek találhatók háromkategóriás Bradley-Terry és Davidson modellek esetén részletes konvergencia-elemzéssel a [15] publikációban. Normális eloszlás használata esetén általában legkisebb négyzetek elvén alapuló paraméterbecslési módszereket alkalmaznak [10].

A maximum likelihood becslések alkalmazása során fontos ellenőrizni, hogy vajon a likelihood függvény felveszi-e a maximumát, és ha igen, egyetlen helyen veszi-e fel. Ha ugyanis a maximumot nem veszi fel a függvény, akkor a becslendő paraméterekre nem kapunk értéket, ha viszont a maximumhely nem egyértelmű, azaz a maximum értéke több helyen is felvétel, akkor vajon melyik helyet tekintjük a becslés értékének. Kétopciós Bradley-Terry modell esetén Ford [14], háromkategóriás modell esetén Davidson [11] adott feltételt a likelihood függvény maximumának létezésére és a maximumhely egyértelműségére. Ezt nagy mértékben általánosították a [16] publikációban. Általános, szigorúan log-konkáv sűrűségfüggvénnyel rendelkező eloszlások esetén a [13] publikációban található elégséges feltétel kettőnél több kategória esetén, ez azonban három kategória esetén gyengébb a [16]-ban található feltételnél.

Mi a páros összehasonlítási modelleket sporteredmények, a labdarúgó Európa-bajnokság eredményeinek kiértékelésére kívántuk használni. A Davidson által adott feltétel, valamint a [13]-beli feltételrendszer nem biztosítja az Európa Bajnokság mérkőzései eredményeinek kiértékelhetőségét, így a [16] publikációban megtalálható feltételrendszert alkalmaztuk és megállapítottuk, hogy ezen általánosabb feltételrendszer alapján sem kiértékelhetők az eredmények 3 kategóriában. Ez vezetett az összetettebb, ötkategóriás modell alkalmazásához. A modellben felállítottunk egy feltételrendszert, ami garantálja a kiértékelhetőséget 5 kategóriában. Ellenőriztük, hogy a mérkőzések eredményei teljesítik

ezt a feltételrendszert, s ez alapján az ötkategóriás modellben kiértékeljük a csapatokat.

A kiértékelések elvégzéséhez numerikus módszert konstruáltunk. A módszer lényege, hogy felírtuk a likelihood függvény logaritmusának parciális deriváltjait, és fixpont-iterációs módszerrel megkerestük a zérushelyet. A kapott argumentum lett a paraméterek maximum likelihood becslése, amely értékek alapján sorba tudtuk állítani a kiértékelendő objektumokat (futballcsapatokat), és becsült erősség-értékeket tudtunk hozzájuk rendelni. A kapott eredményeket összevetettük egymással, valamint más, pontozáson és továbbjutási eredményeken alapuló sorrenddel.

A sztochasztikus háttérű páros összehasonlítási módszerek mellett a páros összehasonlítási módszerek egy másik, gyakran alkalmazott csoportját alkotják a páros összehasonlítási mátrixokon alapuló módszerek. Ezek közül kiemelkedik a Saaty által kidolgozott Analytic Hierarchy Process (AHP) [17]. A módszer rendkívül népszerű, sok ezer publikációban szerepel. Hátránya, hogy kevés mérkőzés esetén a kiértékelés eredménye erősen függ attól, hogy milyen számmal helyettesítjük az egyes kategóriákba eső összehasonlítási eredményeket. Mindazonáltal születtek összehasonlítások az AHP és a sztochasztikus háttérű páros összehasonlítási módszerek által adott eredmények tekintetében, lásd például [18, 19]. Ezek a publikációk a páros összehasonlítási módszerek sportbeli alkalmazásai. A sportágak, amikre alkalmazták az előző publikációkban, a futball és a tenisz volt, de számos egyéb példát is mondhatunk, például a baseballt [20], a sakkot [21], vagy a kézilabdát [22].

Az általunk választott sztochasztikus háttérű páros összehasonlítási háttérű modellek előnye a páros összehasonlítási mátrixokon alapuló módszerekkel szemben az is, hogy alkalmasak előrejelzésre, és végezhetők bennük hipotézisvizsgálatok is. Ugyan ezeket a pozitív tulajdonságokat a mérkőzések kis száma miatt most nem alkalmazzuk, de más esetekben igen hasznosak lehetnek.

A dolgozatban bemutatott eredmények újszerűsége abban rejlik, hogy egyrészt ellenőriztem a kézenfekvőnek tűnő, és a publikációkban futball, kézilabda, sakk, stb. sportágak esetén alkalmazott háromkategóriás modell alkalmazha-

tóságát a 2024-es labdarúgó Európa Bajnokság mérkőzéseinek eredményeire, valamint mivel nemleges választ kaptam, ezért áttértem az ötkategóriás modellre. Ebben a modellben is ellenőriztem az alkalmazhatóság feltételeinek teljesülését, és kidolgoztam egy fixpont iteráción alapuló módszert a numerikus optimalizálásra. Számítógépes programokat készítettem a numerikus számításokhoz és a feltételek ellenőrzéséhez. A kiértékelés eredményei alapján olyan erősségi sorrendet tudtam felállítani a csapatok között, amelyek nem csak a mérkőzések eredményeit veszik figyelembe, hanem azt is, hogy ki ellen érték el azt az eredményt.

A munka eredményei alkalmazhatók más esetekben is, akár sportbajnokságok kiértékelésére, akár más területen, páros összehasonlítások alapján kapott adatok kiértékelésére.

2. fejezet

Sztochasztikus háttérű páros összehasonlítási modellek

A munkám során a 2024. évi labdarúgó Európa-bajnokságon résztvevő labdarúgó nemzeti csapatok teljesítményét fogom összehasonlítani a tornán lejátszott mérkőzések eredményeinek felhasználásával, oly módon, hogy a mérkőzések eredményeit páros összehasonlítások kimeneteleként értelmezem. Összesen 24 csapat játszott 51 meccset. A nemzeti csapatok a tornán belül legfeljebb csak egyszer tudtak egymással játszani.

A páros összehasonlítási módszerek egyik gyakran használt osztálya a sztochasztikus háttérű modellek. Ezen modellekben az összehasonlító objektumok mögé látens valószínűségi változókat képzelünk. Az összehasonlítás eredménye attól függ, hogy a látens valószínűségi változók különbsége mekkora. A látens valószínűségi változók különbségének eloszlása sokféle lehet. Amennyiben normális eloszlásúnak feltételezzük őket, akkor Thurstone modellről beszélünk, amennyiben logisztikus eloszlásúnak, akkor Bradley-Terry modellről. Attól függően, hogy a döntés eredményeként hány kategória megengedett, 2, 3, 5, ..., s -kategóriás modellekkel állunk szemben. 2 kategóriás modellekben csak "jobb" és "rosszabb" döntés lehetséges, míg a több kategóriásban megengedett "egyforma", esetleg "sokkal jobb", "sokkal rosszabb" döntési kategória is.

2.1. Az általánosított Bradley-Terry modell

A dolgozatban az előbb is említett Bradley-Terry modellt fogjuk használni, mely a valószínűségi változók különbségét logisztikus eloszlásúként kezeli. A

paraméterek becslésére maximum likelihood becslést (MLE) alkalmazunk, ami a likelihood függvény vagy a loglikelihood függvény maximalizálásával kapható meg. Mivel mindkét függvény nagyon összetett, így az optimalizálás csak numerikusan végezhető el.

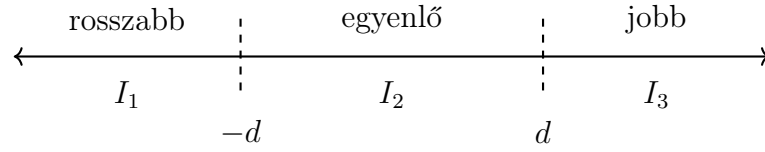
A maximumhely meghatározására fixpont iterációs módszert alkalmazunk. A módszer lényege, hogy megkeressük a parciális deriváltak zérushelyét. A zérushely megkeresése érdekében a parciális deriváltakat egyenlővé tesszük nullával, és kifejezzük belőle a változókat. Ezt a fixpontegyenletet iterációvá írjuk át, amit végrehajtva megkapjuk a szélsőérték helyet.

Mivel labdarúgó Európa-bajnokság eredményeit szeretnénk kiértékelni, ezért a döntetleneket is megengedjük, vagyis a legegyszerűbb modell, amit használhatunk a 3 kategóriás modell. Ennek a "kis vereségek/győzelmek" és "nagy vereségek/győzelmek" különböző értelmezésére alkalmas változata az 5 kategóriás modell. Meg fogjuk mutatni, hogy az eredmények a 3 kategóriás modellben nem kiértékelhetők, ezért a több információt kezelő modellt fogjuk alkalmazni. Nézzük meg részletesen az általunk használt 3 és 5 kategóriás modelleket.

2.1.1. 3 kategóriás modell

A 3 kategóriás modellben a "rosszabb", "egyenlő" és "jobb" döntési lehetőségeket engedjük meg.

Jelöljük a kiértékelendő objektumokat az $1, 2, \dots, n$ számokkal. A mi esetünkben ezek az egyes országok csapatai. A csapatok mögött meghúzódó látens valószínűségi változókat $\xi_1, \xi_2, \dots, \xi_n$ jelölje. Ezek különbsége az $m_i - m_j + \eta_{ij}$, ahol m_i, m_j jelölik az i, j objektumok várható értékét, η_{ij} pedig egy 0 várható értékű valószínűségi változó F eloszlásfüggvényével. Az összehasonlítás eredménye a különbség értéke alapján alakul ki.



2.1. ábra. A döntési kategóriák és a hozzájuk tartozó intervallumok a 3 kategóriás modell esetén.

A 2.1. ábra szerinti intervallumokat a d paraméterrel határoljuk. Ha a különbség kisebb mint $-d$, akkor i "rosszabb", mint j . Ha a különbség a $[-d; d]$ intervallumba esik, akkor a döntés "egyenlő". Ha a különbség meghaladja a d -t, akkor i "jobb", mint j .

Ezen döntések valószínűsége a következőképpen számolható.

$$\begin{aligned} p_{i,j,1}^{(3)} &= P(\xi_i - \xi_j \in I_1) = F(-d - (m_i - m_j)), \\ p_{i,j,2}^{(3)} &= P(\xi_i - \xi_j \in I_2) = F(d - (m_i - m_j)) - F(-d - (m_i - m_j)), \\ p_{i,j,3}^{(3)} &= P(\xi_i - \xi_j \in I_3) = 1 - F(d - (m_i - m_j)). \end{aligned}$$

Ha a Bradley-Terry modellel állunk szemben, akkor az $F(x)$ eloszlásfüggvény

$$F(x) = \frac{1}{1 + e^{-x}}, \quad x \in \mathbb{R}.$$

Definiáljuk az alábbi paramétereket:

$$\begin{aligned} \pi_i &= \frac{e^{m_i}}{\sum_{k=1}^n e^{m_k}}, \\ \theta &= e^d, \theta > 1. \end{aligned}$$

Látható, hogy

(F1) feltétel $0 < \pi_i < 1$ minden $i = 1, \dots, n$ -re, $\sum_{i=1}^n \pi_i = 1$ és $\theta > 1$.

Jelölje az adatokat tartalmazó általánosított mátrixot az $A \in \mathbb{R}^{n \times n \times 3}$. Az $A_{i,j,k}$ mátrixelem az a szám, ahányszor az i és j objektumok összehasonlítása során a k döntés született. Amennyiben $k = 1$ akkor " i rosszabb mint j ", ha

$k = 2$ akkor " i egyenlő j -vel", ha $k = 3$ akkor " i jobb mint j " döntés lett meghozva. Ha $i = j$, akkor legyen $A_{i,j,1} = A_{i,j,2} = A_{i,j,3} = 0$.

Természetesen a döntések szimmetriája miatt ha i jobb mint j , akkor j rosszabb mint i , azaz $A_{i,j,3} = A_{j,i,1}$, illetve $A_{i,j,2} = A_{j,i,2}$.

Algebrai átalakítások után beláthatók az alábbiak:

$$\begin{aligned} p_{i,j,1}^{(3)} &= \frac{\pi_j}{\pi_i\theta + \pi_j}, \\ p_{i,j,2}^{(3)} &= \frac{\pi_i\pi_j(\theta^2 - 1)}{(\pi_i + \pi_j\theta)(\pi_i\theta + \pi_j)}, \\ p_{i,j,3}^{(3)} &= \frac{\pi_i}{\pi_i + \pi_j\theta}. \end{aligned}$$

A likelihood függvény annak a valószínűsége, hogy az A adathalmaz alakul ki, amennyiben a paraméterek $\pi_1, \dots, \pi_n, \theta$. Ennek a valószínűsége (független döntéseket feltételezve):

$$\begin{aligned} L(A|\pi_1, \dots, \pi_n, \theta) &= \prod_{i=1}^{n-1} \prod_{j=i+1}^n \left[\left(\frac{\pi_j}{\pi_i\theta + \pi_j} \right)^{A_{i,j,1}} \right. \\ &\quad \cdot \left(\frac{\pi_i\pi_j(\theta^2 - 1)}{(\pi_i + \pi_j\theta)(\pi_i\theta + \pi_j)} \right)^{A_{i,j,2}} \cdot \left. \left(\frac{\pi_i}{\pi_i + \pi_j\theta} \right)^{A_{i,j,3}} \right]. \end{aligned} \quad (1.1)$$

Az 1.1. függvény logaritmusai:

$$\begin{aligned} \ln(L(A|\pi_1, \dots, \pi_n, \theta)) &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left[A_{i,j,1} \cdot \ln \left(\frac{\pi_j}{\pi_i\theta + \pi_j} \right) \right. \\ &\quad \left. + A_{i,j,2} \cdot \ln \left(\frac{\pi_i\pi_j(\theta^2 - 1)}{(\pi_i + \pi_j\theta)(\pi_i\theta + \pi_j)} \right) + A_{i,j,3} \cdot \ln \left(\frac{\pi_i}{\pi_i + \pi_j\theta} \right) \right]. \end{aligned}$$

Ezt a loglikelihood függvényt tudjuk egyszerűsíteni, így a parciális deriváltjait

könnyebben meg lehet határozni:

$$\begin{aligned} \ln(L(A|\pi_1, \dots, \pi_n, \theta)) = & \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left[[\ln(\pi_j) - \ln(\pi_i\theta + \pi_j)] \cdot A_{i,j,1} \right. \\ & + [\ln(\pi_i) + \ln(\pi_j) + \ln(\theta - 1) + \ln(\theta + 1) - \ln(\pi_i + \pi_j\theta) - \ln(\pi_i\theta + \pi_j)] \cdot A_{i,j,2} \\ & \left. + [\ln(\pi_i) - \ln(\pi_i + \pi_j\theta)] \cdot A_{i,j,3} \right]. \end{aligned} \quad (1.2)$$

A továbbiakban legyen

$$LL = \ln(L(A|\pi_1, \dots, \pi_n, \theta)).$$

A paraméterek maximum likelihood becslése az a $(\hat{\pi}_1, \hat{\pi}_2, \dots, \hat{\pi}_n, \hat{\theta})$ paraméteregyüttes, ami a 1.2. függvényt maximalizálja az (F1) feltételek mellett. A maximumhely megkeresését fixpont iterációval végezzük el. A loglikelihood függvény parciális deriváltjai az alábbi alakban felírhatók:

π_i szerinti parciális derivált (egyenlővé téve 0-val):

$$\begin{aligned} \frac{\partial LL}{\partial \pi_i} = & \sum_{j=i+1}^n \left[(A_{i,j,3} + A_{i,j,2}) \cdot \frac{1}{\pi_i} - (A_{i,j,1} + A_{i,j,2}) \cdot \frac{\theta}{\pi_i\theta + \pi_j} \right. \\ & \left. - (A_{i,j,3} + A_{i,j,2}) \cdot \frac{1}{\pi_i + \pi_j\theta} \right] + \sum_{j=1}^{i-1} \left[(A_{j,i,1} + A_{j,i,2}) \cdot \frac{1}{\pi_i} \right. \\ & \left. - (A_{j,i,1} + A_{j,i,2}) \cdot \frac{1}{\pi_j + \pi_i\theta} - (A_{j,i,3} + A_{j,i,2}) \cdot \frac{\theta}{\pi_j\theta + \pi_i} \right] = 0 \end{aligned}$$

θ szerinti parciális derivált (egyenlővé téve 0-val):

$$\begin{aligned} \frac{\partial LL}{\partial \theta} = & \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left[A_{i,j,2} \cdot \frac{2\theta}{\theta^2 - 1} - (A_{i,j,1} + A_{i,j,2}) \cdot \frac{\pi_i}{\pi_i\theta + \pi_j} \right. \\ & \left. - (A_{i,j,3} + A_{i,j,2}) \cdot \frac{\pi_j}{\pi_i + \pi_j\theta} \right] = 0 \end{aligned}$$

A parciális deriváltak zérushelyének meghatározásához iteratív módszert

használunk, még hozzá a fixpont iterációt:

$$\underline{x}^{(t+1)} = f(\underline{x}^{(t)}).$$

Ha vesszük a loglikelihood függvény π_i és θ szerinti deriváltjait, egyenlővé tesszük őket nullával és átrendezzük az összefüggést, akkor az alábbi iteratív képleteket kapjuk:

$$\pi_i^{(t+1)} = \frac{\sum_{j=1}^{i-1} [A_{i,j,1} + A_{i,j,2}] + \sum_{j=i+1}^n [A_{i,j,2} + A_{i,j,3}]}{P}, \text{ ahol}$$

$$P = \sum_{j=1}^{i-1} \left[(A_{j,i,1} + A_{j,i,2}) \cdot \frac{1}{\pi_j^{(t+1)} + \pi_i^{(t)} \theta^{(t)}} + (A_{j,i,2} + A_{j,i,3}) \cdot \frac{\theta^{(t)}}{\pi_j^{(t+1)} \theta^{(t)} + \pi_i^{(t)}} \right]$$

$$+ \sum_{j=i+1}^n \left[(A_{i,j,1} + A_{i,j,2}) \cdot \frac{\theta^{(t)}}{\pi_j^{(t)} + \pi_i^{(t)} \theta^{(t)}} + (A_{i,j,2} + A_{i,j,3}) \cdot \frac{1}{\pi_j^{(t)} \theta^{(t)} + \pi_i^{(t)}} \right]$$

$$\theta^{(t+1)} = 1 + T \cdot \frac{\sum_{j \neq i} A_{i,j,2}}{N}, \text{ ahol}$$

$$T = \frac{2\theta^{(t)}}{\theta^{(t)} + 1}$$

$$N = \sum_{j \neq i} \left[(A_{i,j,2} + A_{i,j,3}) \cdot \frac{\pi_j^{(t+1)}}{\pi_i^{(t+1)} + \pi_j^{(t+1)} \theta^{(t)}} + \right.$$

$$\left. (A_{i,j,1} + A_{i,j,2}) \cdot \frac{\pi_i^{(t+1)}}{\pi_i^{(t+1)} \theta^{(t)} + \pi_j^{(t+1)}} \right]$$

Ahogy megfigyelhető az iteratív összefüggésekből, a keresés során felhasználjuk az újonnan számított értékeket. A "sima" fixpont iterációhoz képest ezzel egy gyorsabban konvergáló eljárást kapunk.

A maximumkeresésnél az alábbi megállási feltételt alkalmazzuk:

- Az egymást követő iterációk során kapott vektorok különbségének l^2 normáját vizsgálva, $\sqrt{\sum_{i=1}^n (\pi_i^{(t)} - \pi_i^{(t-1)})^2 + (\theta^{(t)} - \theta^{(t-1)})^2} < \varepsilon_1$,

- a parciális deriváltak értéke szerint, azaz, ha egyetlen egy parciális derivált abszolút értéke sem halad meg egy bizonyos értéket, így $\left| \frac{\partial LL}{\partial \pi_i} \right| < \varepsilon_2$ és $\left| \frac{\partial LL}{\partial \theta} \right| < \varepsilon_2$,
- az iterációs szám alapján, azaz, ha egy bizonyos értéknél több iterációt hajtanánk végre; t iterációs szám mellett $t > \frac{1}{\varepsilon_3}$.

Amennyiben az előbb felsorolt megállási feltételek közül legalább egy teljesül, akkor leállítjuk a maximumhely keresését. Ezeknek az $\varepsilon_1, \varepsilon_2, \varepsilon_3$ paramétereknek a változtatásával lehet befolyásolni, hogy mennyire hamar álljon le a keresés. A kevés iterációt végrehajtó keresések előnye a gyors végrehajtás, hátrányuk viszont a nagyobb numerikus hiba.

2.1.2. 5 kategóriás modell

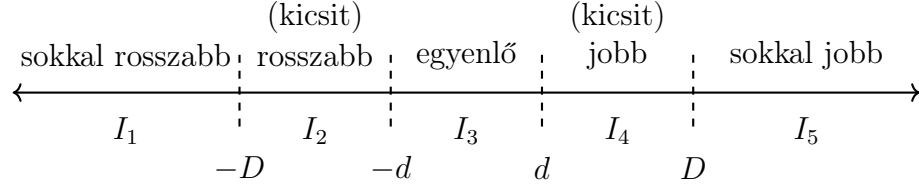
Az 5 kategóriás modell a "rosszabb", "egyenlő" és "jobb" eredménnyel záruló döntéseken felül megenged egy "sokkal rosszabb" és egy "sokkal jobb" döntési kategóriát.

Vannak sporteredmények között olyanok, amik "nagy" győzelemnek és vereségnek, vagy döntetlennek számítanak, és vannak olyanok, amelyekre a "nagy" jelző nem illik rá, ezeket "kicsi" győzelmeknek és vereségnek hívjuk. Ekkor az 5 döntési kategóriának 5 darab intervallum felel meg a számegyenesen. Ezek kijelölésére újabb paraméter bevezetésére van szükségünk.

A 2.2. ábra szerinti intervallumokat a d és D paraméterekkel határoljuk. Általánosított esetben a d érték jelöli az "egyenlő" felső, illetve a "kicsit jobb" döntési kategóriák alsó határát. A D érték jelöli a "kicsit jobb" felső, illetve a "sokkal jobb" döntési kategória alsó határát. Ebből kifolyólag teljesülnie kell a $d < D$ egyenlőtlenségnek.

A szimmetria miatt így ki lesznek jelölve a "kicsit rosszabb/jobb" és a "sokkal rosszabb/jobb" kategóriák határai is. Továbbra is a $\xi_i - \xi_j = m_i - m_j - \eta_{ij}$

különbséggel dolgozunk. Így az alábbiakat mondhatjuk az egyes kategóriákba eső döntések valószínűségeiről:



2.2. ábra. A döntési kategóriák és a hozzájuk tartozó intervallumok az 5 kategóriás modell esetén.

$$\begin{aligned}
 p_{i,j,1}^{(5)} &= P(\xi_i - \xi_j \in I_1) = F(-D - (m_i - m_j)), \\
 p_{i,j,2}^{(5)} &= P(\xi_i - \xi_j \in I_2) = F(-d - (m_i - m_j)) - F(-D - (m_i - m_j)), \\
 p_{i,j,3}^{(5)} &= P(\xi_i - \xi_j \in I_3) = F(d - (m_i - m_j)) - F(-d - (m_i - m_j)), \\
 p_{i,j,4}^{(5)} &= P(\xi_i - \xi_j \in I_4) = F(D - (m_i - m_j)) - F(d - (m_i - m_j)), \\
 p_{i,j,5}^{(5)} &= P(\xi_i - \xi_j \in I_5) = 1 - F(D - (m_i - m_j)).
 \end{aligned}$$

Az összehasonlítások eredményeit az $A \in \mathbb{R}^{n \times n \times 5}$ általánosított mátrixban gyűjtjük, ahol $A_{i,j,k}$, $k = 1 \dots 5$ az a szám, ahány esetben az i és j objektumok összehasonlítása során a k döntés született. Amennyiben $k = 1$ akkor "sokkal rosszabb", ha $k = 2$ akkor "kicsit rosszabb", ha $k = 3$ akkor "egyenlő", ha $k = 4$ akkor "kicsit jobb", végül pedig $k = 5$ esetében "sokkal jobb" döntés lett meghozva. A döntések szimmetriája miatt ha i sokkal jobb, mint j , akkor j sokkal rosszabb, mint i , azaz $A_{i,j,5} = A_{j,i,1}$, $A_{i,j,4} = A_{j,i,2}$, illetve $A_{i,j,3} = A_{j,i,3}$. Most is igaz, hogy $A_{i,i,k} = 0$, $k = 1, 2, 3, 4, 5$.

Bevezetjük az alábbi paramétereket:

$$\begin{aligned}\pi_i &= \frac{e^{m_i}}{\sum_{k=1}^n e^{m_k}}, \\ \theta &= e^d, \\ \beta &= e^D.\end{aligned}$$

Most is teljesül, hogy

(F2) feltétel $0 < \pi_i < 1$ minden $i = 1, 2, \dots, n$ -re, $\sum_{i=1}^n \pi_i = 1$, $1 < \theta < \beta$.

Algebrai átalakítások után beláthatók az alábbiak:

$$\begin{aligned}p_{i,j,1}^{(5)} &= \frac{\pi_j}{\pi_i \beta + \pi_j}, \\ p_{i,j,2}^{(5)} &= \frac{\pi_i \pi_j (\beta - \theta)}{(\pi_i \beta + \pi_j) (\pi_i \theta + \pi_j)}, \\ p_{i,j,3}^{(5)} &= \frac{\pi_i \pi_j (\theta^2 - 1)}{(\pi_i + \pi_j \theta) (\pi_i \theta + \pi_j)}, \\ p_{i,j,4}^{(5)} &= \frac{\pi_i \pi_j (\beta - \theta)}{(\pi_i + \pi_j \beta) (\pi_i + \pi_j \theta)}, \\ p_{i,j,5}^{(5)} &= \frac{\pi_i}{\pi_i + \pi_j \beta}.\end{aligned}$$

A likelihood függvény a következőképpen alakul:

$$\begin{aligned}L(A|\pi_1, \dots, \pi_n, \theta, \beta) &= \prod_{i=1}^{n-1} \prod_{j=i+1}^n \left[\left(\frac{\pi_j}{\pi_i \beta + \pi_j} \right)^{A_{i,j,1}} \right. \\ &\cdot \left(\frac{\pi_i \pi_j (\beta - \theta)}{(\pi_i \beta + \pi_j) (\pi_i \theta + \pi_j)} \right)^{A_{i,j,2}} \cdot \left(\frac{\pi_i \pi_j (\theta^2 - 1)}{(\pi_i + \pi_j \theta) (\pi_i \theta + \pi_j)} \right)^{A_{i,j,3}} \\ &\cdot \left(\frac{\pi_i \pi_j (\beta - \theta)}{(\pi_i + \pi_j \beta) (\pi_i + \pi_j \theta)} \right)^{A_{i,j,4}} \cdot \left. \left(\frac{\pi_i}{\pi_i + \pi_j \beta} \right)^{A_{i,j,5}} \right] \quad (2.1)\end{aligned}$$

A 2.1. függvény logaritmusa:

$$\begin{aligned}
 \ln(L(A|\pi_1, \dots, \pi_n, \theta, \beta)) &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left[A_{i,j,1} \cdot \ln \left(\frac{\pi_j}{\pi_i \beta + \pi_j} \right) \right. \\
 &+ A_{i,j,2} \cdot \ln \left(\frac{\pi_i \pi_j (\beta - \theta)}{(\pi_i \beta + \pi_j) (\pi_i \theta + \pi_j)} \right) + A_{i,j,3} \cdot \ln \left(\frac{\pi_i \pi_j (\theta^2 - 1)}{(\pi_i + \pi_j \theta) (\pi_i \theta + \pi_j)} \right) \\
 &\left. + A_{i,j,4} \cdot \ln \left(\frac{\pi_i \pi_j (\beta - \theta)}{(\pi_i + \pi_j \beta) (\pi_i + \pi_j \theta)} \right) + A_{i,j,5} \cdot \ln \left(\frac{\pi_i}{\pi_i + \pi_j \beta} \right) \right]. \quad (2.2)
 \end{aligned}$$

A paraméterek maximum likelihood becslése az a $(\hat{\pi}_1, \hat{\pi}_2, \dots, \hat{\pi}_n, \hat{\theta}, \hat{\beta})$ paraméteregyüttes, ami a 2.2. függvényt maximalizálja az (F2) feltételek mellett. A maximumhely keresést ismét fixpont iterációval végezzük el. A loglikelihood függvény az alábbi alakban is írható:

$$\begin{aligned}
 \ln(L(A|\pi_1, \dots, \pi_n), \theta, \beta) &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left[[\ln(\pi_j) - \ln(\pi_i \beta + \pi_j)] \cdot A_{i,j,1} \right. \\
 &+ [\ln(\pi_i) + \ln(\pi_j) + \ln(\beta - \theta) - \ln(\pi_i \beta + \pi_j) - \ln(\pi_i \theta + \pi_j)] \cdot A_{i,j,2} \\
 &+ [\ln(\pi_i) + \ln(\pi_j) + \ln(\theta - 1) + \ln(\theta + 1) - \ln(\pi_i + \pi_j \theta) - \ln(\pi_i \theta + \pi_j)] \cdot A_{i,j,3} \\
 &+ [\ln(\pi_i) + \ln(\pi_j) + \ln(\beta - \theta) - \ln(\pi_i + \pi_j \beta) - \ln(\pi_i + \pi_j \theta)] \cdot A_{i,j,4} \\
 &\left. + [\ln(\pi_i) - \ln(\pi_i + \pi_j \beta)] \cdot A_{i,j,5} \right].
 \end{aligned}$$

A továbbiakban legyen:

$$LL^{(5)} = \ln(L(A|\pi_1, \dots, \pi_n, \theta, \beta)).$$

A loglikelihood függvény parciális deriváltjai (egyenlővé téve 0-val):

$$\begin{aligned}
 \frac{\partial LL^{(5)}}{\partial \pi_i} &= \sum_{j=1}^{i-1} \left[(A_{j,i,1} + A_{j,i,2} + A_{j,i,3} + A_{j,i,4}) \cdot \frac{1}{\pi_i} \right. \\
 &\quad - (A_{j,i,4} + A_{j,i,5}) \cdot \frac{\beta}{\pi_i \beta + \pi_j} - (A_{j,i,1} + A_{j,i,2}) \cdot \frac{1}{\pi_i + \pi_j \beta} \\
 &\quad \left. - (A_{j,i,3} + A_{j,i,4}) \cdot \frac{\theta}{\pi_i \theta + \pi_j} - (A_{j,i,2} + A_{j,i,3}) \cdot \frac{1}{\pi_i + \pi_j \theta} \right] \\
 &+ \sum_{j=i+1}^n \left[(A_{i,j,2} + A_{i,j,3} + A_{i,j,4} + A_{i,j,5}) \cdot \frac{1}{\pi_i} - (A_{i,j,1} + A_{i,j,2}) \cdot \frac{\beta}{\pi_i \beta + \pi_j} \right. \\
 &\quad - (A_{i,j,4} + A_{i,j,5}) \cdot \frac{1}{\pi_i + \pi_j \beta} - (A_{i,j,3} + A_{i,j,2}) \cdot \frac{\theta}{\pi_i \theta + \pi_j} \\
 &\quad \left. - (A_{i,j,3} + A_{i,j,4}) \cdot \frac{1}{\pi_i + \pi_j \theta} \right] = 0,
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial LL^{(5)}}{\partial \theta} &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n (A_{i,j,2} + A_{i,j,4}) \cdot \frac{-1}{\beta - \theta} + A_{i,j,3} \cdot \frac{2\theta}{\theta^2 - 1} \\
 &- (A_{i,j,2} + A_{i,j,3}) \cdot \frac{\pi_i}{\pi_i \theta + \pi_j} - (A_{i,j,3} + A_{i,j,4}) \cdot \frac{\pi_j}{\pi_i + \pi_j \theta} = 0,
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial LL^{(5)}}{\partial \beta} &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n (A_{i,j,2} + A_{i,j,4}) \cdot \frac{1}{\beta - \theta} \\
 &- (A_{i,j,1} + A_{i,j,2}) \cdot \frac{\pi_i}{\pi_i \beta + \pi_j} - (A_{i,j,4} + A_{i,j,5}) \cdot \frac{\pi_j}{\pi_i + \pi_j \beta} = 0.
 \end{aligned}$$

Iteratív képletek a fentiekből kialakítva:

$$\begin{aligned}
 \pi_i^{(t+1)} &= \frac{1}{P} \cdot \left[\sum_{j=1}^{i-1} (A_{i,j,1} + A_{i,j,2} + A_{i,j,3} + A_{i,j,4}) \right. \\
 &\quad \left. + \sum_{j=i+1}^n (A_{i,j,2} + A_{i,j,3} + A_{i,j,4} + A_{i,j,5}) \right], \text{ ahol} \\
 P &= \sum_{j=1}^{i-1} \left[\frac{1}{\pi_i^{(t)} \beta^{(t)} + \pi_j^{(t+1)}} \cdot (A_{j,i,1} + A_{j,i,2}) + \frac{\beta^{(t)}}{\pi_i^{(t)} + \pi_j^{(t+1)} \beta^{(t)}} \cdot (A_{j,i,4} + A_{j,i,5}) \right. \\
 &\quad \left. + \frac{1}{\pi_i^{(t)} \theta^{(t)} + \pi_j^{(t+1)}} \cdot (A_{j,i,2} + A_{j,i,3}) + \frac{\theta^{(t)}}{\pi_i^{(t)} + \pi_j^{(t+1)} \theta^{(t)}} \cdot (A_{j,i,3} + A_{j,i,4}) \right] \\
 &\quad + \sum_{j=i+1}^n \left[\frac{\beta^{(t)}}{\pi_i^{(t)} \beta^{(t)} + \pi_j^{(t)}} \cdot (A_{j,i,1} + A_{j,i,2}) + \frac{1}{\pi_i^{(t)} + \pi_j^{(t)} \beta^{(t)}} \cdot (A_{j,i,4} + A_{j,i,5}) \right. \\
 &\quad \left. + \frac{\theta^{(t)}}{\pi_i^{(t)} \theta^{(t)} + \pi_j^{(t)}} \cdot (A_{j,i,2} + A_{j,i,3}) + \frac{1}{\pi_i^{(t)} + \pi_j^{(t)} \theta^{(t)}} \cdot (A_{j,i,3} + A_{j,i,4}) \right], \\
 \theta^{(t+1)} &= 1 + \frac{2\theta^{(t)}}{\theta^{(t)} + 1} \cdot \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n A_{i,j,3}}{T}, \text{ ahol} \\
 T &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left[\frac{\pi_i^{(t+1)}}{\pi_i^{(t+1)} \theta^{(t)} + \pi_j^{(t+1)}} \cdot (A_{i,j,2} + A_{i,j,3}) \right. \\
 &\quad \left. + \frac{\pi_j^{(t+1)}}{\pi_i^{(t+1)} + \pi_j^{(t+1)} \theta^{(t)}} \cdot (A_{i,j,3} + A_{i,j,4}) + \frac{1}{\beta^{(t)} - \theta^{(t)}} \cdot (A_{i,j,2} + A_{i,j,4}) \right], \\
 \beta^{(t+1)} &= \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n [A_{i,j,2} + A_{i,j,4}]}{B} + \theta^{(t+1)}, \text{ ahol} \\
 B &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left[\frac{\pi_i^{(t+1)}}{\pi_i^{(t+1)} \beta^{(t)} + \pi_j^{(t+1)}} \cdot (A_{i,j,1} + A_{i,j,2}) \right. \\
 &\quad \left. + \frac{\pi_j^{(t+1)}}{\pi_i^{(t+1)} + \pi_j^{(t+1)} \beta^{(t)}} \cdot (A_{i,j,4} + A_{i,j,5}) \right].
 \end{aligned}$$

A megállási feltételeknél a 3 kategóriás modellben említetteket használjuk, az-
 zal a kiegészítéssel, hogy a β paraméterre is vizsgáljuk a feltételeket.

Míg a háromopciós modell numerikus megoldásával részletesen foglalkoztak [15], az öttopciós modellel kapcsolatban nem találtam eredményeket az irodalomban. A fent említett iteratív módszer konvergenciájának elméleti vizsgálata további kutatásokat igényel.

3. fejezet

Kiértékelhetőségi feltételrendszerek

A likelihood, illetve a loglikelihood függvény maximalizálásakor szembekerülünk azzal a kérdéssel, hogy vajon létezik-e a maximum, és ha létezik, akkor vajon egyértelmű-e. Ha nem létezik a maximum, akkor nem kapunk becült értékeket. Ha nem egyértelmű a maximumhely, akkor nem tudjuk, hogy melyik értékeket tekintsük becült értékeknek. Így a létezés és egyértelműség vizsgálata nem csak elméleti kérdés, hanem a kiértékelhetőség elvégzésének elemi feltétele.

A maximum létezését és egyértelműségét a 3 és 5 kategóriás modellekben más feltételrendszerekkel vizsgálhatjuk.

3.1. Feltételrendszer 3 kategóriás modell esetén

A kiértékelhetőségre 3 kategóriás modell esetén az irodalomban megtalálható feltételrendszer az alábbi [16]:

SC 1. Létezik legalább egy indexpár (i_1, j_1) , amelyekre $0 < A_{i,j,2}$ teljesül. Ezt az jelenti, hogy kell lennie legalább egy döntetlen eredménynek.

Megjegyezzük, hogy ha ez nem teljesül, akkor belátható, hogy a loglikelihood függvény szigorúan monoton nő, ha d tart a 0-hoz. Ekkor a 2 kategóriás modellre kell áttérni.

SC 2. Az összehasonlítások eredményei segítségével definiáljunk egy DG irányított gráfot, amelynek a csúcsai az egyes összehasonlítandó objektumokat képviselik. Amennyiben $0 < A_{i,j,3}$ teljesül, az i csúcsból egy j csúcsba vezető

irányított élt húzzunk be. Amennyiben $0 < A_{i,j,1}$ teljesül, a j csúcsból az i csúcsba vezető irányított élt húzzunk. A DG gráfban léteznie kell egy körnek.

SC 3. Bármely nem üres S és $\bar{S}$ objektumokat tartalmazó partíciókra létezik kettő, nem feltétlenül különböző (i_2, j_2) és (i_3, j_3) , $i_2, i_3 \in S$, $j_2, j_3 \in \bar{S}$ indexpár, melyekre $0 < A_{i_2, j_2, 3}$ és $0 < A_{i_3, j_3, 1}$ teljesül, vagy létezik olyan (i_4, j_4) , $i_4 \in S$, $j_4 \in \bar{S}$ indexpár, amelyre $0 < A_{i_4, j_4, 2}$ teljesül.

Vagyis bárhogyan osztjuk 2 részre az objektumokat, mindig találhatunk a részhalmazok között "egyenlő" döntést, vagy oda-vissza "jobb" és "rosszabb" döntést.

Megjegyezzük, hogy ez a feltétel szükséges a kiértékelhetőséghez, mert ha nem teljesül, akkor az egyik részhalmaz minden eleme jobb, mint a másik részhalmaz bármely eleme, de nem lehet megmondani, hogy mennyivel. A likelihood függvény egy bizonyos irányban szigorúan monoton nőni fog.

Ezen feltételek közül az egyik teljesülése nem garantálja bármely másinak a teljesülését.

Az SC 1., SC 2. és SC 3. feltételek együttesen elégségesek a kiértékelhetőséghez, de nem szükségesek. Van olyan eset, amikor kiértékelhetők az adatok, de nem teljesül az SC 2. feltétel.

A feltételek ellenőrzése számítógépes program segítségével elvégezhető. Számítástechnikai szempontból az SC 1. feltétel a legkedvezőbb, mivel egyetlen "egyenlő" él elég a teljesüléséhez, így $O(|E|)$ időkomplexitással rendelkezik, ahol E jelöli a DG irányított gráf éleinek halmazát.

Az SC 2. feltétel ellenőrzésének műveletigénye $O(|V| + |E|)$, ahol V jelöli a DG irányított gráf csúcsainak halmazát, amelyek egyben az összehasonlítandó objektumok.

Az SC 3. feltétel ellenőrzése érdekében vegyük figyelembe az alábbiakat:

SC 3. módosítása Definiáljunk egy FDG irányított gráfot, amelynek a csúcsai az egyes összehasonlítandó objektumokat képviselik. Amennyiben $0 <$

$A_{i,j,3}$ teljesül, az i csúcsból egy j csúcsba vezető él van. Amennyiben $0 < A_{i,j,1}$ teljesül, a j csúcsból az i csúcsba vezető él van. Amennyiben $0 < A_{i,j,2}$ teljesül, az i és j csúcs között egy kétirányú él van. Ennek az FDG gráfnak erősen összefüggőnek kell lennie.

Az eredeti SC 3. feltétel és az előbb említett módosítása ekvivalensek. Az irányított gráf erős összefüggőség-ellenőrzésre léteznek $\mathcal{O}(|V| + |E|)$ gyorsaságú algoritmusok, ahol V jelöli az FDG gráf csúcsainak halmazát E pedig az éleinek halmazát.

Az összefüggőség ellenőrzését gyorsíthatjuk az alábbi megfontolásokkal. Defináljunk egy G egyszerű gráfot, amelynek a csúcsai az egyes összehasonlítható objektumokat képviselik. Amennyiben tetszőleges $i, j, i \neq j$ csúcsokra $0 < \sum_{k=1}^3 A_{i,j,k}$ teljesül, akkor az i és j csúcs között van pontosan egy irányítatlan él. Jelölje E a G gráf éleit, V pedig a csúcsait. Az adatok kiértékelhetőségéhez szükséges az alábbi két feltétel:

4. G összefüggő.

5. G élszámára teljesül, hogy $|E| \geq |V| - 1$.

6. Észrevehető továbbá, hogy amennyiben létezik egy tetszőleges i csúcs, melynek a t csúcsszámára teljesül a $t = |V| - 1$ feltétel, akkor a gráf összefüggő.

Ezzel a G gráffal az élek számából is lehet következtetni a gráf összefüggőségére. Amennyiben $|E| < |V| - 1$, a G gráf nem összefüggő. Amennyiben létezik egy csúcs, amelynek csúcsszáma $t = |V| - 1$, akkor a gráf összefüggő. Amennyiben 4 és 5 fennáll de 6 nem, akkor egy szélességi bejárás algoritmus-sal érdemes megvizsgálni a gráf összefüggőségét. A G gráf összefüggősége azt bizonyítja, hogy nincsen olyan objektum vagy objektumhalmaz, amely nem volt összehasonlítva más objektumokkal vagy objektumhalmazokkal. A G gráf összefüggőségéből nem következik sem SC 1., sem SC 2., sem SC 3., amennyiben viszont G nem összefüggő, az SC 3. feltétel nem teljesül.

A végrehajtás gyorsítása érdekében először érdemes vizsgálni a G gráf összefüggőségét, majd utána érdemes az SC 1., az SC 2. és az SC 3. feltételeket a felsorolás szerinti sorrendben ellenőrizni.

3.2. Feltételrendszer 5 kategóriás modell esetén

A 3 kategóriás modell kiértékelhetőségi feltételrendszereinek vizsgálata alapján az alábbi feltételeket mondhatjuk ki 5 kategóriás modell esetén:

SC 1. Van olyan (i_1, j_1) indexpár, amelyre $A_{i_1, j_1, 1} > 0$, (i_2, j_2) amelyre $A_{i_2, j_2, 2}$, illetve (i_3, j_3) amelyre $A_{i_3, j_3, 3} > 0$ teljesül.

Szóban kifejezve ez azt jelenti, hogy minden kategóriába esik döntés.

SC 2. Definiáljunk egy DG irányított gráf, amelynek a csúcsai az egyes összehasonlítandó objektumokat képviselik. Amennyiben $0 < A_{i, j, 4} + A_{i, j, 5}$ teljesül, az i csúcsból egy j csúcsba vezető irányított él van. Amennyiben $0 < A_{i, j, 1} + A_{i, j, 2}$ teljesül, a j csúcsból az i csúcsba vezető irányított él van. Ebben a DG gráfban kell lennie egy olyan körnek, amelyre teljesül az, hogy létezik a körben olyan (i_6, j_6) indexpár, melyre $0 < A_{i_6, j_6, 5}$.

SC 3. Bármely nem üres S és $\bar{S}$ partíciókra létezik kettő, nem feltétlenül különböző (i_7, j_7) és (i_8, j_8) , $i_7, i_8 \in S$, $j_7, j_8 \in \bar{S}$ indexpár, melyekre $0 < A_{i_7, j_7, 5}$ és $0 < A_{i_8, j_8, 1}$ teljesül, vagy létezik olyan (i_9, j_9) , $i_9 \in S$, $j_9 \in \bar{S}$ indexpár, amelyre $0 < \sum_{k=2}^4 A_{i_9, j_9, k}$ teljesül.

Ez azt jelenti, hogy bármely két részhalmaz össze van kötve vagy közbülső döntéssel, vagy oda-vissza "sokkal jobb" döntéssel.

Most is megállapíthatjuk, hogy ezen feltételek közül az egyik teljesülése nem garantálja bármely másiknak a teljesülését. Számítástechnikai szempontból az SC 1. feltétel a legkedvezőbb, mivel egyetlen "döntetlen", "egyszerű vereség" vagy "nagy vereség" él elég a teljesüléséhez, így $O(|E|)$ időkomplexitással rendelkezik, ahol E jelöli a DG irányított gráf éleinek halmazát.

Most is igaz, hogy az SC 2. feltétel ellenőrzésének műveletigénye $\mathcal{O}(|V| + |E|)$, ahol V jelöli a DG irányított gráf csúcsainak halmazát.

SC 3. módosítása Definiáljunk egy FDG irányított gráfot, amelynek a csúcsai az egyes összehasonlítandó objektumokat képviselik. Amennyiben $0 < A_{i,j,5}$ teljesül, az i csúcsból egy j csúcsba vezető él van. Amennyiben $0 < A_{i,j,1}$ teljesül, a j csúcsból az i csúcsba vezető él van. Amennyiben $0 < \sum_{k=2}^4 A_{i,j,k}$ teljesül, az i és j csúcs között egy kétirányú él van. Ennek az FDG gráfnak erősen összefüggőnek kell lennie.

Az eredeti SC 3. feltétel és az előbb említett módosítása ekvivalensek. Az irányított gráf erős összefüggőségének ellenőrzésére léteznek $\mathcal{O}(|V| + |E|)$ gyorsaságú algoritmusok, ahol V jelöli az FDG gráf csúcsainak halmazát, E pedig az éleinek halmazát.

Most is gyorsíthatjuk a feltételek ellenőrzését az alábbi észrevételekkel: Definiáljunk egy G egyszerű gráfot, amelynek a csúcsai az egyes összehasonlítandó objektumok. Amennyiben tetszőleges $i, j, i \neq j$ csúcsokra $0 < \sum_{k=1}^5 A_{i,j,k}$ teljesül, akkor az i és j csúcs között van pontosan egy irányítatlan él. Jelölje E a G gráf éleit, V pedig a csúcsait. A kiértékelhetőség szükséges feltételei az alábbiak:

4. G összefüggő.

5. G élszámára teljesül, hogy $|E| \geq |V| - 1$.

6. Megállapíthatjuk még, hogy amennyiben létezik egy tetszőleges i csúcs, melynek a t csúcsszámára teljesül a $t = |V| - 1$ feltétel, akkor a G gráf összefüggő.

A felsoroltakat figyelembe véve egy programot készítettem a kiértékelhetőség ellenőrzésére mind a 3, mind az 5 kategóriás modellben.

4. fejezet

A 2024. évi UEFA labdarúgó Európa-bajnokság eredményeinek kiértékelése

A 2024. évi UEFA labdarúgó-bajnokságon összesen 24 nemzeti csapat vett részt. A nemzeti csapatok négyes csoportokra osztva játszottak körmérkőzéses rendszerben egymással az egyenes kieséses szakaszba kerülésért. Az egyenes kieséses szakaszba kerülést a szerzett pontszámok, illetve holtverseny esetén másodlagos mutatók alapján döntötték el. A győzelemért 3 pont, a döntetlenért mindkét csapatnak 1-1 pont, a vereségért 0 pont járt. A kieséses szakaszban a rendes játékidő végén döntetlen eredmény esetén hosszabbítás következett. Ha ekkor is döntetlen lett az eredmény, akkor 11-es rúgásokkal döntötték el, hogy ki a továbbjutó.

4.1. A mérkőzések eredményei

A bajnokság során először csoportonként 6 meccs lejátszásával dőlt el, hogy ki jut tovább a kieséses szakaszba. A kieséses szakaszban, amennyiben a hagyományos, 90 perces játékidőben döntetlen lett az eredmény, akkor egy 30 perces hosszabbítás következett. Amennyiben a hosszabbítás során is döntetlen lett az eredmény, 11-esekkel döntötték el a továbbjutást. Az eredménytáblázatban zárójelekkel külön fel van tüntetve, hogy melyik mérkőzés esetén volt hosszabbítás, illetve mi lett a 11-es rúgások eredménye.

A következőkben az országokat az angol elnevezésük rövidítése azonosítja. A 4.1., 4.2., 4.3., 4.4., 4.5. és 4.6. táblázatokban az egyes csoportkörökben lejátszott mérkőzések eredményeit, a 4.7., 4.8. és 4.9. táblázatok az egyenes

kieséses szakasz mérkőzéseinek eredményeit, a 4.10. táblázat a döntő eredményét láthatjuk. Az eredményeket az UEFA hivatalos oldaláról töltöttük le [23].

Mérkőzés	Eredmény
GER - SCO	5-1
HUN - SUI	1-3
GER - HUN	2-0
SCO - SUI	1-1
SUI - GER	1-1
SCO - HUN	0-1

4.1. táblázat. Az A csoport mérkőzéseinek eredményei.

Mérkőzés	Eredmény
ESP - CRO	3-0
ITA - ALB	2-1
CRO - ALB	2-2
ESP - ITA	1-0
ALB - ESP	0-1
CRO - ITA	1-1

4.2. táblázat. A B csoport mérkőzéseinek eredményei.

Mérkőzés	Eredmény
SLO - DEN	1-1
SER - ENG	0-1
SLO - SER	1-1
DEN - ENG	1-1
ENG - SLO	0-0
DEN - SER	0-0

4.3. táblázat. A C csoport mérkőzéseinek eredményei.

Mérkőzés	Eredmény
POL - NED	1-2
AUS - FRA	0-1
POL - AUS	1-3
NED - FRA	0-0
NED - AUS	2-3
FRA - POL	1-1

4.4. táblázat. A D csoport mérkőzéseinek eredményei.

Mérkőzés	Eredmény
ROM - UKR	3-0
BEL - SVK	0-1
SVK - UKR	1-2
BEL - ROM	2-0
SVK - ROM	1-1
UKR - BEL	0-0

4.5. táblázat. Az E csoport mérkőzéseinek eredményei.

Mérkőzés	Eredmény
TUR - GEO	3-1
POR - CZE	2-1
GEO - CZE	1-1
TUR - POR	0-3
GEO - POR	2-0
CZE - TUR	1-2

4.6. táblázat. Az F csoport mérkőzéseinek eredményei.

Mérkőzés	Eredmény
ESP - GEO	4-1
GER - DEN	1-0
POR - SLO	0-0 (0-0) (3-0)
FRA - BEL	1-0
ROM - NED	0-3
AUS - TUR	1-2
ENG - SVK	1-1 (2-1)
SUI - ITA	2-0

4.7. táblázat. A nyolcaddöntők mérkőzéseinek eredményei.

Mérkőzés	Eredmény
ESP - GER	1-1 (2-1)
POR - FRA	0-0 (0-0) (3-5)
NED - TUR	2-1
ENG - SUI	1-1 (1-1) (5-3)

4.8. táblázat. A negyeddöntők mérkőzéseinek eredményei.

Mérkőzés	Eredmény
ESP - FRA	2-1
NED - ENG	1-2

4.9. táblázat. Az elődöntők mérkőzéseinek eredményei.

Mérkőzés	Eredmény
ESP - ENG	2-1

4.10. táblázat. A döntő eredménye

4.2. Hagyományos kiértékelések eredményei

A ponttáblázatok a győzelmek és döntetlenek számát veszik figyelembe. Ezek alapján rendelnek pontokat az egyes csapatokhoz. Amennyiben a kapott pont-

eredmény alapján holtverseny van, másodlagos mutatók alapján rendezik sorba a csapatokat.

A legtöbb bajnokságban, amely egy csoportkörrel és egy egyenes kieséses szakasszal rendelkezik, csak a csoportkörben lejátszott mérkőzések alapján vezetnek ponttáblázatot. Ez a 2024. évi UEFA labdarúgó Európa-bajnokság esetén is igaz.

A 4.11. táblázat a csoportmérkőzések, illetve az egyenes kieséses szakasz során szerzett pontok mellett a továbbjutások számát is figyelembe veszi. Ha valaki továbbjut a nyolcaddöntőkből, akkor megelőzi azokat, akik nem jutottak tovább. Ugyanez igaz a negyeddöntők, illetve az elődöntők mérkőzései alapján is. Amennyiben holtverseny alakul ki, a jobb gólkülönbség, majd a több szerzett gól, végül a kisebb fair play pontszám alapján vannak rangsorolva a csapatok. A 11-es rúgások utáni győzelem ez esetben döntetlennek számít, és így a meccset játszó mindkét csapat 1-1 pontot kap. Mivel a bronzmérkőzést nem játszották le, ezért az elődöntőkben szereplő francia és holland csapat ugyanannyi továbbjutással rendelkezik. A holland csapat a több győzelme miatt került a franciák elé.

Mivel az UEFA hivatalosan nem rakta sorba az összes csapatot, ezért a nemhivatalos, az előbb részletezett rendezőelvek szerint elkészített 4.11 táblázat eredményeihez hasonlítjuk a páros összehasonlítási módszerek segítségével kapott kiértékelési eredményeket.

A 2024. évi UEFA labdarúgó Európa-bajnokság eredményeinek kiértékelése

Helyezés	Csapat	M	Gy	D	V	G+	G–	GK	Pontszám
1.	ESP	7	7	0	0	15	4	+11	21
2.	ENG	7	3	3	1	8	6	+2	12
3.	NED	6	3	1	2	10	7	+3	10
3.	FRA	6	2	3	1	4	3	+1	9
5.	GER	5	3	1	1	11	4	+7	10
6.	SUI	5	2	3	0	8	4	+4	9
7.	TUR	5	3	0	2	8	8	0	9
8.	POR	5	2	2	1	5	3	+2	8
9.	AUS	4	2	0	2	7	6	+1	6
10.	BEL	4	1	1	2	2	2	0	4
11.	SLO	4	0	4	0	2	2	0	4
12.	SVK	4	1	1	2	4	5	-1	4
13.	ROM	4	1	1	2	4	6	-2	4
14.	ITA	4	1	1	2	3	5	-2	4
15.	GEO	4	1	1	2	5	8	-3	4
16.	DEN	4	0	3	1	2	4	-2	3
17.	UKR	3	1	1	1	2	4	-2	4
18.	HUN	3	1	0	2	2	5	-3	3
19.	SER	3	0	2	1	1	2	-1	2
20.	CRO	3	0	2	1	3	6	-3	2
21.	ALB	3	0	1	2	3	5	-2	1
22.	CZE	3	0	1	2	3	5	-2	1
23.	POL	3	0	1	2	3	6	-3	1
24.	SCO	3	0	1	2	2	7	-5	1

4.11. táblázat. A 2024. évi labdarúgó Európa-bajnokság nemhivatalos pont-táblája.

A következő alfejezetekben megpróbáljuk elvégezni a mérkőzések eredményeinek kiértékelését páros összehasonlítási módszerekkel a 3 és 5 kategóriás modellekben.

4.3. A mérkőzések eredményeinek kiértékelése 3 kategóriás modellben.

Először vizsgáljuk meg, hogy kiértékelhetők-e az eredmények a 3 kategóriás modellben.

A 3 kategóriás modellben a győzelmek "jobb", a vereségek "rosszabb", a döntetlenek "egyenlő" döntéseknek értendők. A kieséses szakaszban, amennyiben a rendes játékidőben döntetlen alakul ki, akkor hosszabbításban döntik el, hogy ki jut tovább. Ezeket a hosszabbítással végződő meccseket értelmezhetjük döntetlenként, győzelemként, vagy ezek valamely súlyozással történő kombinációjaként. A tornán összesen 14 döntetlennel, illetve 37 győzelemmel, vereséggel záródó meccs játszódott le. Amennyiben a rendes játékidő után megnyert mérkőzéseket döntetlennek tekintjük, akkor ebből 19 döntetlen, illetve 32 győzelem lesz.

Gyakran a rendes játékidő után megnyert mérkőzéseket döntetlenként értelmezik, mivel amennyiben nem tud az egyik csapat a rendes játékidő alatt győzni, abból arra lehet következtetni, hogy a két csapat között nincs nagy különbség. Azonban a sportszakemberek szemszögéből ez nem megfelelő értékelés, hiszen a mérkőzés az egyik csapat továbbjutásával végződött. A hosszabbításban megnyert mérkőzések győzelemnek értelmezése esetén lesz egy határozott döntés, mely szerint a továbbjutó csapat jobb, mint a másik. A 3 kategóriás modellben azt is megtehetjük, hogy a hosszabbítás után megnyert mérkőzésekre adunk egy döntetlent, és emellett egy fél győzelmet.

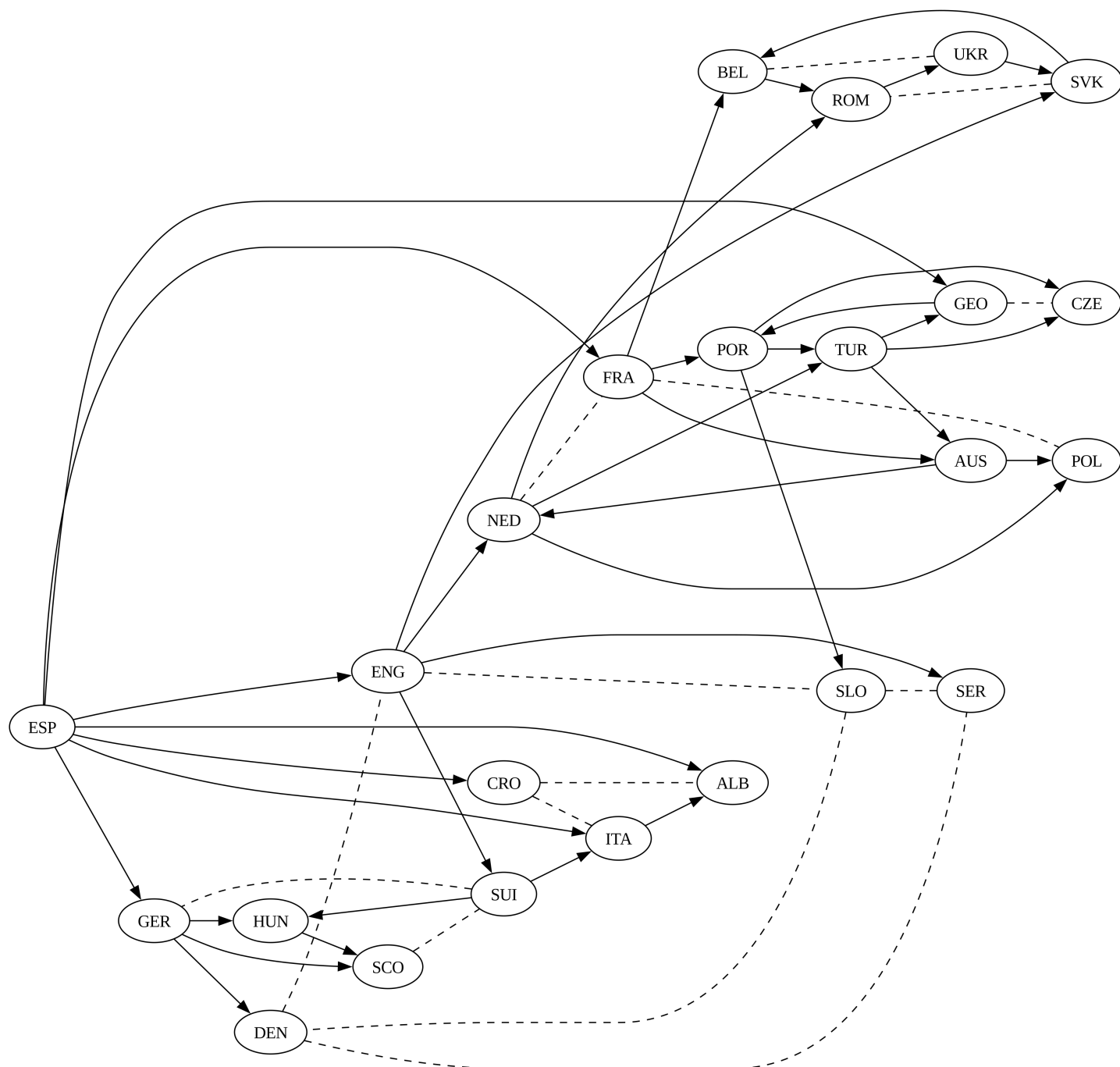
4.3.1. A mérkőzések eredményeinek kiértékelhetősége

A továbbiakban 3 különálló esetet fogunk vizsgálni, melyek szerint a rendes játékidő után elért győzelem:

1. teljesértékű győzelemnek számít (A);
2. teljesértékű döntetlennek számít (B);

3. egy teljesértékű döntetlennek és egy félértékű győzelemnek számít (C).
Mivel a loglikelihood függvényben az $A_{i,j,k}$ szorzótényezők, ezért nem okoz problémát, ha nem egészek.

Az összes esetben ugyanúgy kezeljük a 11-esekkel megnyert mérkőzéseket, mint a hosszabbításban megnyerteket.



4.1. ábra. Az eredményekből származtatott gráf a 3 kategóriás modellben az A eset szerint. ($\longrightarrow$ győzelem, - - - döntetlen)

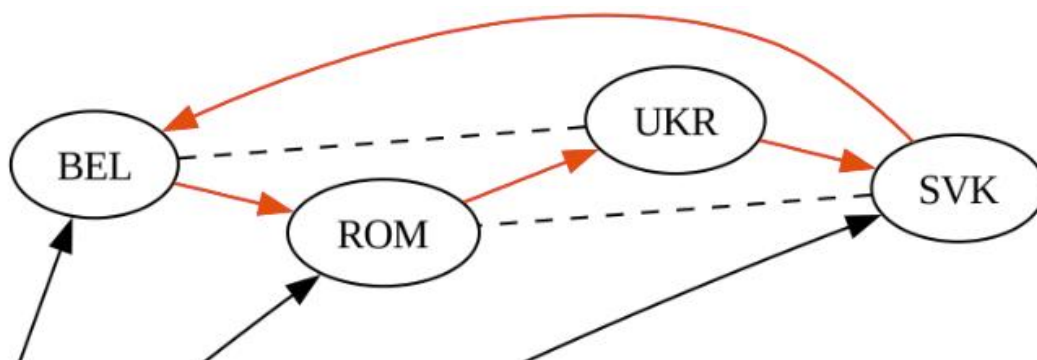
A 4.1. ábra mutatja az összes mérkőzés eredménye alapján felrajzolt gráfot az A eset alapján, ahol szaggatott vonallal a döntetlenek, nyíllal pedig a győ-

zelmek vannak jelölve. Az ábra alapján látszik, hogy ez a gráf összefüggő.

Vizsgáljuk meg, hogy teljesülnek-e a 3 kategóriás modellre vonatkozó SC 1., SC 2. és SC 3. feltételek.

SC 1. teljesül, mert van döntetlen, például a belga-ukrán mérkőzés eredménye.

SC 2. teljesül, mivel van csupa győzelemből álló kör.



4.2. ábra. A 3 kategóriás modellben az A eset szerint kapott csupa győzelmekből álló kör.

SC 3. nem teljesül, mivel összesen 4 olyan partíciót is találtunk, amely nem felel meg a követelményeknek. Ezek a következők:

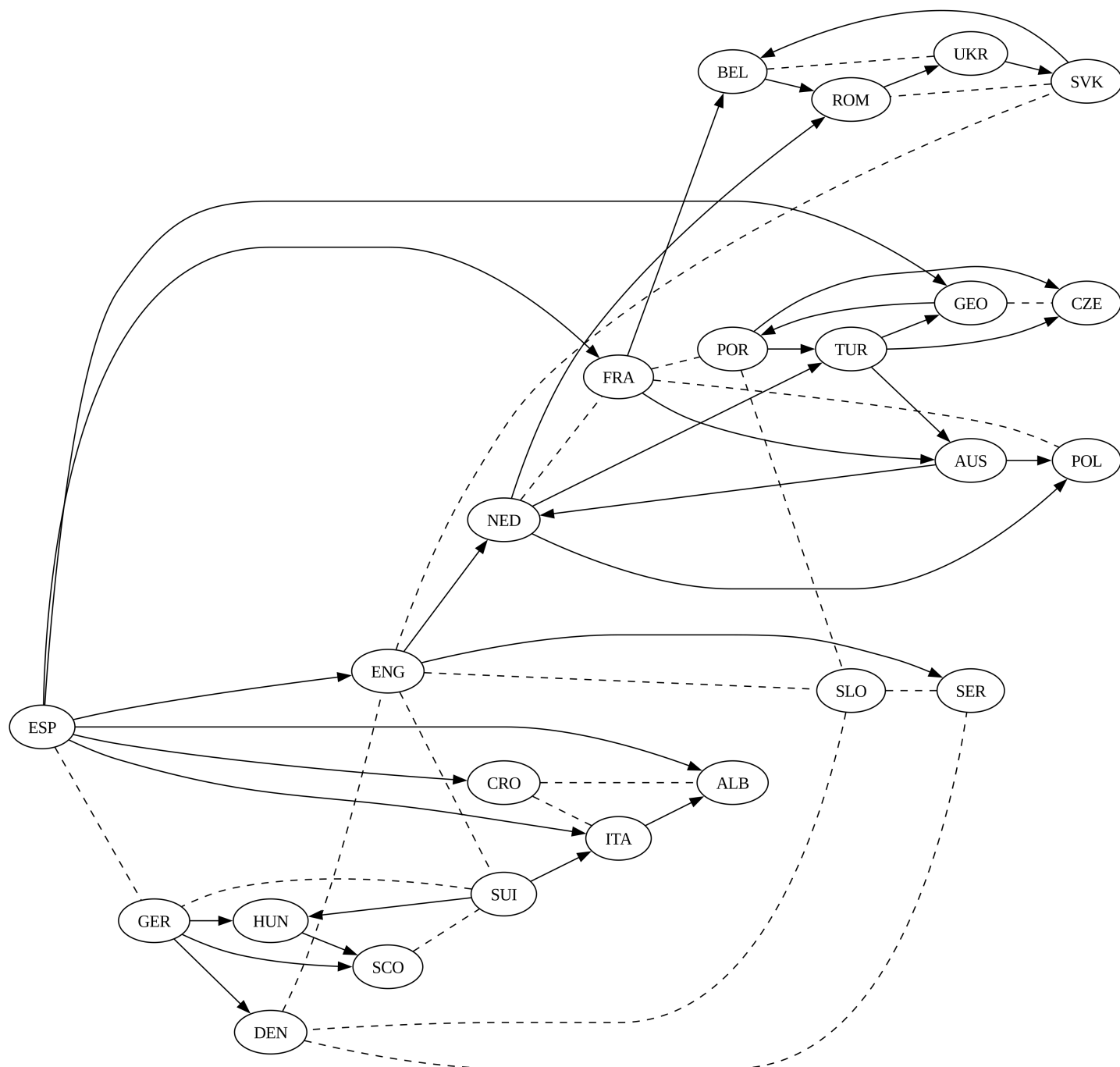
- Spanyolország, szemben az összes többivel;
- Horvátország, Albánia és Olaszország, szemben az összes többivel;
- Belgium, Románia, Ukrajna és Szlovákia, szemben az összes többivel;
- Horvátország, Albánia, Olaszország, Belgium, Románia, Ukrajna és Szlovákia, szemben az összes többivel.

Mivel az SC 3. feltétel teljesülése szükséges a kiértékelhetőséghez, ezért a 3 kategóriás modellben az A esetben kapott eredmények nem kiértékelhetők.

A felsorolt csapatok eltávolításával egy 16 csapatból álló, már 3 kategóriában kiértékelhető részhalmazt kaphatunk. A kiértékelés eredményeit a 4.3.2. alfejezet tartalmazza.

Térjünk rá a B esetre.

A B esetbeli eredmények alapján felrajzolt gráfot a 4.3. ábrán láthatjuk. Ez a gráf más, mint az előző, például a spanyol és a német csapatot egy döntetlen köti össze, szemben az A esettel, amikor az összekötés egy győzelemmel történik.

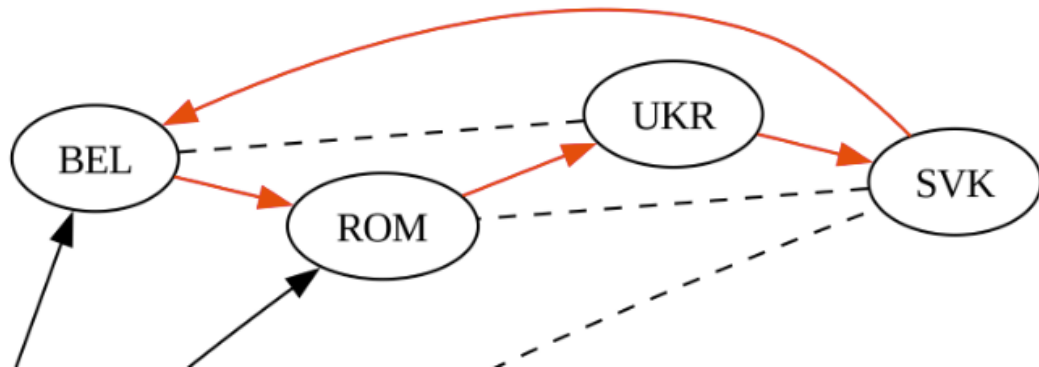


4.3. ábra. A B eset szerinti eredményekből származtatott gráf a 3 kategóriás modellben. ($\longrightarrow$ győzelem, - - - döntetlen)

Vizsgáljuk meg most is az SC 1., SC 2. és SC 3. feltételek teljesülését.

SC 1. teljesül, mert van döntetlen, például a belga-ukrán mérkőzés eredménye.

SC 2. teljesül, mert 4.4. ábrán látható kör csupa győzelmekből álló kör. A körben szereplő 4 ország eredménye ugyanaz, mint az A esetben, de másként kapcsolódnak a többi csapathoz.



4.4. ábra. A csupa győzelmekből álló kör a B eset szerinti eredmények esetén a 3 kategóriás modellben.

SC 3. nem teljesül, mert létezik egy olyan partíció, amely nem felel meg a feltételnek, ez pedig:

- Horvátország, Albánia és Olaszország szemben az összes többivel.

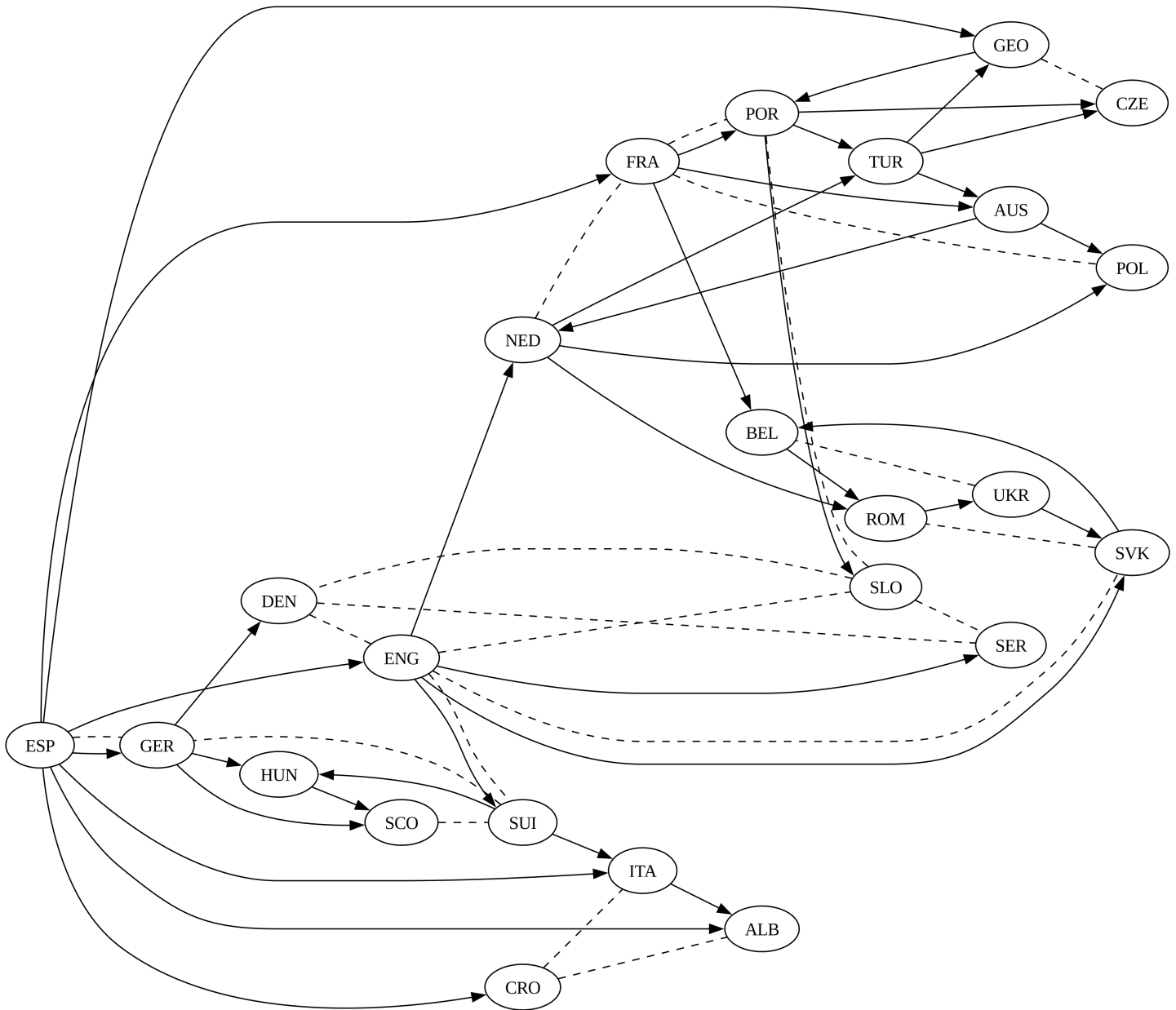
Most kevesebb, a feltételnek nem megfelelő partíciót lehet találni, mivel ebben az esetben például Spanyolország Németországgal egy döntetlennel van összekapcsolva. Az A esetbeli többi partíció az angol-szlovák mérkőzés rendes játékidőben született döntetlen eredménye miatt lett összekötve a többi csapattal.

A horvát, albán és olasz csapatokat kivéve a 24 helyett 21 csapat marad az eredeti mezőnyből, és ez a részhalmaz már kiértékelhető az általuk játszott mérkőzések eredményei alapján.

Térjünk rá a C esetre.

A C esetbeli eredmények alapján a 4.5. ábrán látható gráf a 4.1. és a 4.3 ábrán látható gráfok összeolvasztása. Összesen 5 mérkőzésnél van különbség:

az angol-szlovák, a portugál-szlovén, a spanyol-német, a portugál-francia és az angol-svájci mérkőzéseknél van a győzelmi él mellett egy döntetlen él. Ezeket a meccseket hosszabbításban vagy 11-esek során nyerték meg.



4.5. ábra. A C eset szerinti eredményekből származtatott gráf a 3 kategóriás modellben. (→ győzelem, --- döntetlen)

Most is megállapítható, hogy SC 1. és SC 2. teljesül, de SC 3. nem, pontosan ugyanazon 3 csapat miatt, mint az előző, B esetben.

Megállapíthatjuk, hogy 3 kategóriás modell alkalmazása eredménytelennek bizonyult a 2024. évi labdarúgó Európa-bajnokság adathalmazán akár a A, akár a B, akár a C eset szerint értelmezzük a meccsek eredményét. Általánosságban elmondható, hogy a 3 kategóriás modellel való kiértékelhetetlenség tényéből nem következik az, hogy 5 kategóriás modellel nem lehetne kiértékelhető az adathalmaz [24].

Mivel a 3 kategóriás modellel nem tudjuk kiértékelni a 2024. évi UEFA labdarúgó Európa-bajnokság mérkőzéseinek eredményeit, kénytelenek vagyunk áttérni a több információt felhasználó 5 kategóriás modellre. Ezzel a 4.4. alfejezetben fogunk foglalkozni.

4.3.2. A kiértékelhető részhalmaz eredményei a 3 kategóriás modellben

4.11.	Csapat	Helyezés	$\hat{\pi}_i$	$\hat{m}_i$
1	ESP	—	—	—
2	ENG	2	0,17978	3,46918
3	NED	4	0,07912	2,64839
4	FRA	3	0,16424	3,37875
5	GER	1	0,18155	3,479
6	SUI	9	0,03713	1,89188
7	TUR	6	0,06691	2,48086
8	POR	5	0,07512	2,59649
9	AUS	7	0,05923	2,3588
10	BEL	—	—	—
11	SLO	10	0,03306	1,77569
12	SVK	—	—	—
13	ROM	—	—	—
14	ITA	—	—	—
15	GEO	12	0,02223	1,37888
16	DEN	8	0,0375	1,9017
17	UKR	—	—	—
18	HUN	14	0,01089	0,66536
19	SER	11	0,02491	1,49252
20	CRO	—	—	—
21	ALB	—	—	—
22	CZE	15	0,00716	0,24566
23	POL	13	0,01558	1,02367
24	SCO	16	0,0056	0

4.12. táblázat. Az A eset szerinti kiértékelhető részhalmaz eredményei összevetve a nemhivatalos eredményekkel.

A 4.12. táblázat az A eset szerinti eredmények alapján értékeli azt a 16 csapatot, amely a kiértékelhető részhalmazba tartozik. Az első oszlop a 4.11.

nemhivatalos táblázatbeli sorrendet tartalmazza, a Helyezés nevű oszlop a 3 kategóriás Bradley-Terry modell szerinti sorrendet. A $\hat{\pi}_k$ oszlopban a becsült erősségeket, az utolsó $\hat{m}_i$ oszlopban pedig az objektum mögött meghúzódó valószínűségi változó várható értékét 5 tizedesjegyre kerekített értékeit tüntettük fel. Ezeket a becsült erősségeket akár előrejelzésre is használhatjuk jövőbeli mérkőzések eredményei valószínűségének becslésére. Bár a spanyol együttes mindenkit legyőzött, nyilvánvalóan az első helyre sorolandó, de erősséget nem tudunk hozzárendelni, így előrejelzést sem tudunk tenni vele kapcsolatban.

A német csapat a 16 csapatból álló sor elejére kívánczik, ami azért érdekes, mert ők nem jutottak be a döntőbe. A magyar szövetségi kapitány meg is jegyezte egy riportban, hogy a német csapat a második a legerősebb csapat, csak túl korán találkoztak a spanyolokkal. A páros összehasonlítási módszerrel való kiértékelés alapján kicsivel nagyobb az erősségük, mint az angol csapaté. Végignézve a mérkőzések eredményeit ez indokoltnak tűnik. A spanyol csapaton kívül nem verte meg egyetlen csapat sem őket, de az angoloknak több a döntetlenjük, mint a németeknek. Az angolok a szlovénokkal, a dánokkal, a svájciakkal, illetve a szlovákokkal játszottak döntetlent. A németek viszont csak a svájciakkal jutottak döntetlenre, a többi csapatot megverték. Mind a dán, mind a szlovák, mind a szlovén csapat a középmezőnyben található, így az ellenük játszott döntetlen eredmény az angol csapat erősségét lefele mozdítja.

A magyar csapat a harmadik harmadba került. Ez egybevág azzal, hogy nem sikerült továbbjutniuk a csoportkörből.

4.11.	Csapat	Helyezés	$\hat{\pi}_i$	$\hat{m}_i$
1	ESP	1	0,51376	5,21615
2	ENG	6	0,0377	2,60396
3	NED	4	0,04075	2,68177
4	FRA	3	0,08397	3,40489
5	GER	2	0,08787	3,45022
6	SUI	7	0,03614	2,56184
7	TUR	8	0,03184	2,4352
8	POR	5	0,04023	2,66908
9	AUS	9	0,02782	2,30017
10	BEL	18	0,00531	0,64435
11	SLO	10	0,02108	2,02275
12	SVK	12	0,01007	1,28427
13	ROM	19	0,00504	0,5925
14	ITA	—	—	—
15	GEO	14	0,00889	1,15932
16	DEN	11	0,01406	1,61783
17	UKR	17	0,00606	0,7765
18	HUN	15	0,00733	0,96651
19	SER	13	0,00923	1,19724
20	CRO	—	—	—
21	ALB	—	—	—
22	CZE	21	0,00279	0
23	POL	16	0,00651	0,84709
24	SCO	20	0,00355	0,2403

4.13. táblázat. A B eset szerinti kiértékelhető részhalmaz eredményei összevetve a nemhivatalos eredményekkel.

A 4.13. táblázatban találhatjuk a B eset szerinti eredmények alapján történő kiértékelést, amelyben 21 csapatot tudtunk rangsorolni.

A 21 csapat között már megtalálható a spanyol csapat, amely az első helyre került. A német csapat lett a második helyezett, erősségét tekintve (0,0878) messze lemaradva a spanyoloktól (0,5137). Az angolok visszacsúsznak a ha-

totdik helyre, mivel a spanyol csapat ellen volt egy vereségük. Erősségük körülbelül fele a németekének. Az A esetben ugyan szintén a németek mögött végeztek, most azonban nagyobb a különbség az erősségük között. A mostani nagyobb különbséget az indokolja, hogy a németek döntetleneztek a B eset szerinti kiértékelésben a spanyolokkal, míg az angolok kikaptak tőlük. A belga csapat a tizennyolcadik helyezést érte el, mivel mind a franciák, mind a szlovákok megverték őket. Ezek mindketten előkelőbb helyen szerepelnek a listán, a tizenhetedik helyezett ukránokkal játszottak egy döntetlent, valamint a harmatgyenge román csapatot sikerült megverniük. A belga csapat gyenge teljesítményét alátámasztja a két veresége, és egy gyenge csapattal szemben elért döntetlenje, valamint egy gyenge csapattal szembeni győzelme. Mivel azonban a gyűjtött pontok elegendőek voltak a továbbjutáshoz, ezért a nemhivatalos sorrend alapján tizedik helyre kerültek. Ez is mutatja, hogy a páros összehasonlítási módszerekkel történő kiértékelések realisabb eredményt adnak.

A magyar csapat helyzete hasonló az előbbihez: az utolsó harmadban helyezkedik el, kissé megelőzve Lengyelországot és némileg lemaradva Grúziától/Georgiától.

4.11.	Csapat	Helyezés	$\hat{\pi}_i$	$\hat{m}_i$
1	ESP	1	0,64271	6,39272
2	ENG	4	0,02917	3,3003
3	NED	6	0,01458	2,60648
4	FRA	5	0,02135	2,98829
5	GER	2	0,17248	5,07732
6	SUI	3	0,03011	3,332
7	TUR	8	0,01139	2,35978
8	POR	7	0,01424	2,58344
9	AUS	11	0,00872	2,09249
10	BEL	20	0,00191	0,57155
11	SLO	10	0,00994	2,22312
12	SVK	15	0,00396	1,30453
13	ROM	19	0,00191	0,57637
14	ITA	—	—	—
15	GEO	16	0,0035	1,17959
16	DEN	9	0,01031	2,26038
17	UKR	17	0,00225	0,73674
18	HUN	12	0,00839	2,05379
19	SER	13	0,00594	1,70797
20	CRO	—	—	—
21	ALB	—	—	—
22	CZE	21	0,00108	0
23	POL	18	0,00208	0,65771
24	SCO	14	0,00399	1,31022

4.14. táblázat. A C eset szerinti kiértékelhető részhalmaz eredményei összevetve a nemhivatalos eredményekkel.

A 4.14. táblázatban láthatjuk a C eset szerinti kiértékelés eredményeit, ahol a B esethez hasonlóan 21 csapat van. Mivel itt a hosszabbítás után nyert mérkőzések mind döntetlenként, mind győzelemként figyelembe vannak véve, ezért sorrendmódosulás alakul ki mind az A, mind a B esethez képest.

Most is az első a spanyol csapat óriási előnnyel a második helyezett német csapathoz képest. Ebben az esetben a leginkább szembetűnő változás a svájci csapat előretörése. Ez azért van, mert a svájciak döntetlent értek el a németek ellen a csoportkörben, illetve az angol csapat ellen csak hosszabbításban kaptak ki. Az angolok itt az A esethez képest gyengébbnek számítanak, mert a szlovák csapatot is csak hosszabbítás után sikerült legyőzniük. A belga csapat az előzőekhez képest még rosszabb helyre került, mivel az őket megverő szlovákok erősebbnek számítanak az angolok ellen hosszabbításban elszenvedett vereség miatt.

A különböző esetek egymással és a 4.11. ponttáblázattal való összefüggőségét több korrelációs együttható kiszámításával fogjuk elemezni. Mivel az összehasonlításokban a legkisebb elemszámú csoport az A eset 16 csapattal, ezért mindegyik más csoport létszámát 16 csapatra fogjuk csökkenteni, elhagyva Spanyolországot, Horvátországot, Albániát, Olaszországot, Belgiumot, Romániát, Ukrajnát és Szlovákiát.

4.3.3. Egyes kiértékelési eredmények közötti korrelációk

Az első korrelációs együttható a Pearson korrelációs együttható, amely a becsült erősségek alapján tud számítani összefüggőségi mértéket. Amennyiben van két, x és y vektorunk $\bar{x}$ és $\bar{y}$ átlagokkal, akkor a Pearson korrelációs együttható képlete a következő [25]:

$$r_{x,y} = \frac{\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (K1)$$

	A	B	C
A	1	0,55459	0,68294
B	0,55459	1	0,98191
C	0,68294	0,98191	1

4.15. táblázat. A 3 kategóriás kiértékelési eredmények erősségeinek összevetése a Pearson-féle korrelációs együtthatót használva.

A 4.15. táblázatban láthatjuk a 3 eset közötti Pearson-féle korrelációs együtthatókat. Ezekből megállapíthatjuk, hogy a B és a C esetek között erős az összefüggés 0,98 értékkel, az A és a C esetek már kevésbé összefüggők 0,68 értékkel, míg az A és B esetek gyengén kapcsolódnak össze 0,55 értékkel. Mivel a 4.11. táblázat nem rendelkezik csapatokhoz rendelt erősségekkel, így azt a sorrendet ki kellett hagynunk ebből az összehasonlításból.

A következő korrelációs együttható, amelyet használni fogunk, a Kendall-féle rangkorreláció. Ez sorrendek összefüggőségét méri. A Kendall-féle rangkorrelációt a következőképpen tudjuk kiszámítani [26]:

$$\tau_{x,y} = 1 - \frac{2U}{\binom{n}{2}}, \quad (\text{K2})$$

ahol U jelöli a diszkordáns párok számát. Legyen x_i az i -edik objektum sorban elfoglalt helye egy szempont szerint, y_i pedig egy másik szempont szerint. Akkor tekintjük diszkordánsnak az (x_i, y_i) és (x_j, y_j) párokat, ha $x_i < x_j$ és $y_i > y_j$, vagy $x_i > x_j$ és $y_i < y_j$ teljesül.

	A	B	C	4.11.
A	1	0,3	0,61667	0,75
B	0,3	1	0,68333	0,25
C	0,61667	0,68333	1	0,56667
4.11.	0,75	0,25	0,56667	1

4.16. táblázat. A 3 kategóriás kiértékelési eredmények rangsorainak összevetése egymással, a nemhivatalos 4.11. táblázat rangsorával, Kendall rangkorrelációt használva.

A Kendall-féle rangkorreláció kiszámított együtthatói a 4.16. táblázatban találhatók. Megfigyelhető, hogy itt kifejezetten gyenge az összefüggés az A és B eset között (0,3). A B és C esetek közötti összefüggés inkább közepes, mint erős (0,68), és az A és C esetek közötti összefüggés is közepesnek tekinthető (0,61). A ponttáblához való viszonyulás esetenként nagyon változó: az A eset viszonylag jól korrelál hozzá (0,75), a B eset nagyon gyengén (0,25), míg a C

eset közepesen gyengén (0,56). Az A eset és a nemhivatalos ponttábla sorrendje közötti korrelációérték azzal indokolható, hogy mind az A eset, mind a ponttáblázat a továbbjutásokat jutalmazza.

Megvizsgáljuk még a Spearman-féle rangkorrelációs együtthatókat is. Ez az együttható jellegzetesen abszolút értékben 1-hez közelebbi értékeket ad, mint a Kendall-féle rangkorreláció, annak ellenére, hogy ugyanazokat az adatokat dolgozza fel. Az alábbiak szerint lehet kiszámolni [25]:

$$r_{x,y} = 1 - \frac{6 \sum_{i=1}^n (x_i - y_i)^2}{n(n^2 - 1)}. \quad (\text{K3})$$

	A	B	C	4.11.
A	1	0,33824	0,77059	0,90882
B	0,33824	1	0,81176	0,34412
C	0,77059	0,81176	1	0,75588
4.11.	0,90882	0,34412	0,75588	1

4.17. táblázat. A 3 kategóriás kiértékelési eredmények rangsorainak összevetése egymással és a nemhivatalos 4.11. táblázat rangsorával, Spearman rangkorrelációt használva.

A 4.17. táblázatban találhatók a Spearman-féle rangkorreláció kiszámított értékei. Itt már szorosabban függ össze a B és a C eset (0,81), az A és B esetek között most is nagy az eltérés (0,33), az A és a C között pedig közepes összefüggés van (0,77). Az A eset meglepően közel áll a 4.11. táblázat rangsorához (0,91), a C közepesen függ össze vele (0,77), míg a B eset erősen eltér tőle (0,34).

Ebben az esetben is lényegében ugyanazt a jelenséget láthatjuk, mint a Kendall-féle rangkorrelációnál. Megállapíthatjuk, hogy a 3 kategóriában kiértékelhető részhalmaz eredményei erősen függenek a hosszabbításban eldölt mérkőzések eredményeinek értelmezésétől.

4.4. A mérkőzések eredményeinek kiértékelése 5 kategóriás modell alkalmazása esetén

Az 5 kategóriás modell használata során el kell döntenünk, hogy mely eredményeket tekintjük nagy győzelemnek, kis győzelemnek, kis vereségnek és nagy vereségnek. Ezt úgy is meg lehet határozni, hogy egy olyan gól- vagy pontkülönbséget állítunk be a végeredménynél, amely ha egy mérkőzés során kialakul, akkor az esetek többségében már nem lehet megváltoztatni a mérkőzés kimenetelét. Ez a határ sportáganként lehet különböző. A mi esetünkben labdarúgást vizsgálunk, így a 3 vagy nagyobb gólkülönbséggel való győzelmet hívjuk nagy győzelemnek, az 1 vagy 2 gólkülönbséggel végződő mérkőzéseket pedig kis győzelemnek, kis vereségnek tekintjük. Természetesen ezt másként is megtehetnénk. A 2024. évi labdarúgó Európa-bajnokságnak sem változott meg egyetlen mérkőzésének eredménye, amikor kialakult egy 3 gólkülönbséges vezetés.

Amennyiben ezt a 3 vagy nagyobb gólkülönbség-határt választjuk, akkor az általunk feldolgozandó adathalmazban összesen 31 kis győzelem, illetve kis vereség, 6 nagy győzelem, illetve nagy vereség, és 14 döntetlen lesz. Amennyiben a rendes játékidő után megnyert mérkőzések eredményét döntetleneknek tekintjük, úgy 19-re nő a döntetlenek száma, kis győzelmekből pedig 26 lesz, a nagy győzelmek száma nem változik.

4.4.1. Eredmények különböző súlyozások alkalmazása esetén

Az 5 kategóriás modell használata esetén is van lehetőségünk többféleképpen értelmezni a kieséses szakaszban, rendes játékidőben döntetlennel, hosszabbításban győzelemmel vagy vereséggel végződő mérkőzéseket. Itt is tudjuk használni az A, a B és a C eseteket. A meglévő esetek mellé bevezetünk 2 új esetet is, ezek a következők:

- A rendes játékidőben megnyert mérkőzések 1 győzelemnek, a hosszabbításban megnyert mérkőzések 1 döntetlennek és 0,5 győzelemnek, a 11-es rúgások során megnyert mérkőzések pedig 1,4 döntetlennek és 0,1 győ-

zelemnek számítanak (D). Ebben az esetben tehát egy hosszabbításban megnyert mérkőzés 1,5 súllyal szerepel a loglikelihood függvényben.

- A rendes játékidőben megnyert mérkőzések 1 győzelemnek, a hosszabbításban megnyert mérkőzések 1:0,5 arányban, a 11-es rúgások során megnyert mérkőzések pedig 1,4:0,1 arányban számítanak döntetlennek és győzelemnek (E). Ebben az esetben egy hosszabbításban megnyert mérkőzés 1 súllyal szerepel a loglikelihood függvényben.

Ezekkel az esetekkel lehetőségünk van külön kezelni a hosszabbításban és a 11-es rúgásokban megnyert mérkőzéseket. A D eset annyiban hasonlít a C esethez, hogy a hosszabbításban és a 11-esekkel megnyert mérkőzések fajsúlyosabbak a rendes játékidőben megnyert mérkőzéseknél, mivel összesen 1,5 egységnyi eredményként számítanak. Az E esetben a rendes játékidőben megnyert mérkőzések ugyanakkora fajsúllyal rendelkeznek, mint a rendes játékidőn túl megnyert mérkőzések. A C, a D és az E esetek kiértékelhetőség szempontjából ekvivalensek.

Először térjünk rá az A eset kiértékelhetőségének vizsgálatára.

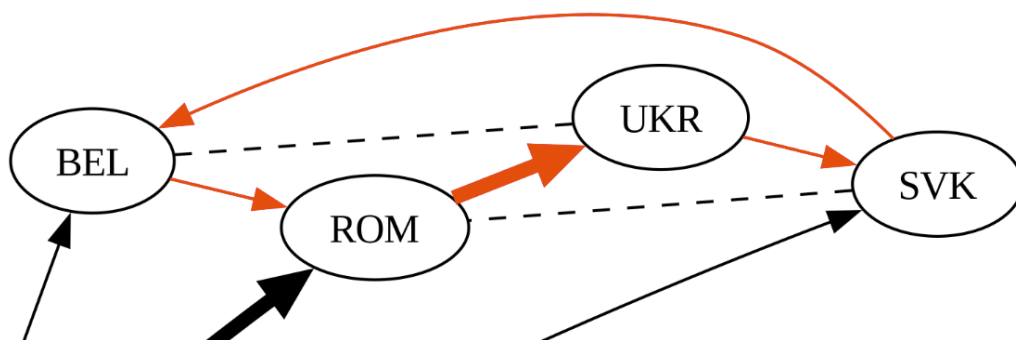
A 4.7. ábra mutatja az összes mérkőzés eredménye alapján felrajzolt gráfot az A esetben. A vastag irányított élek jelölik a nagy győzelmeket, a vékony irányított élek jelölik a kis győzelmeket, a szaggatott irányítatlan vonalak pedig a döntetleneket.

A gráfról könnyen megállapíthatjuk, hogy összefüggő. Vizsgáljuk meg az SC 1., SC 2. és SC 3. feltételeket, melyeket a 3.2. alfejezetben definiáltunk.

Az SC 1. feltétel teljesül, mivel létezik nagy vereség (spanyol-horvát mérkőzés), kis vereség (spanyol-német mérkőzés), illetve döntetlen (horvát-albán mérkőzés) is. Természetesen a kis vereség a másik csapat szempontjából kis győzelem, a nagy vereség pedig nagy győzelem a döntések szimmetriája végett.

Az SC 2. feltétel teljesül, mivel létezik egy olyan kör, amely csak győzel-

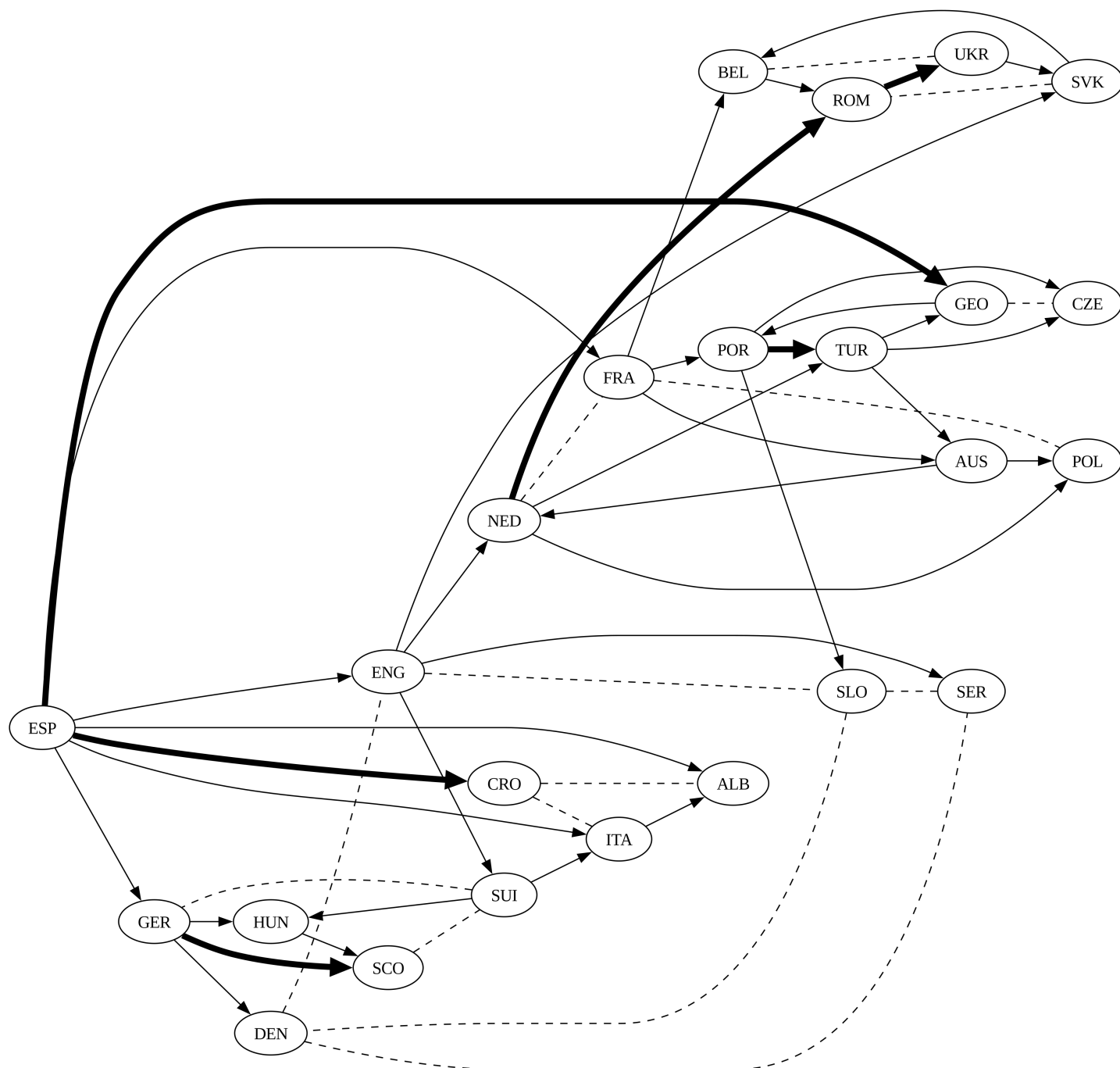
mezből áll, és tartalmaz egy nagy győzelmet, lásd 4.6. ábra.



4.6. ábra. A csupa győzelmekből álló kör az A eset szerinti eredmények az 5 kategóriás modellben.

Az SC 3. feltétel teljesül, mivel minden partíció között van vagy köztes döntés, vagy oda-vissza nagy győzelem/vereség páros. Ezt számítógépes program segítségével ellenőriztük.

Az SC 1., SC 2. és SC 3. feltételek teljesülése miatt az A eset szerinti eredmények alapján kiértékelhetők a csapatok.



4.7. ábra. Az A eset szerint kapott eredmények alapján felrajzolt gráf az 5 kategóriás modellben. ($\Rightarrow$ nagy győzelem, $\rightarrow$ kis győzelem, - - - döntetlen)

A kiértékelések eredményei a 4.18 táblázatban találhatók. Éppen úgy, mint a 3 kategóriás esetben, az első oszlop a 4.11. nemhivatalos táblázat szerinti sor-

rendet rendel a második oszlopban található nemzeti csapatokhoz. A Helyezés nevű oszlop a kiértékelés eredményeként kapott sorrend, a $\hat{\pi}_i$ és az $\hat{m}_i$ oszlopok pedig tartalmazzák az erősségeket, illetve az objektumok várható értékét.

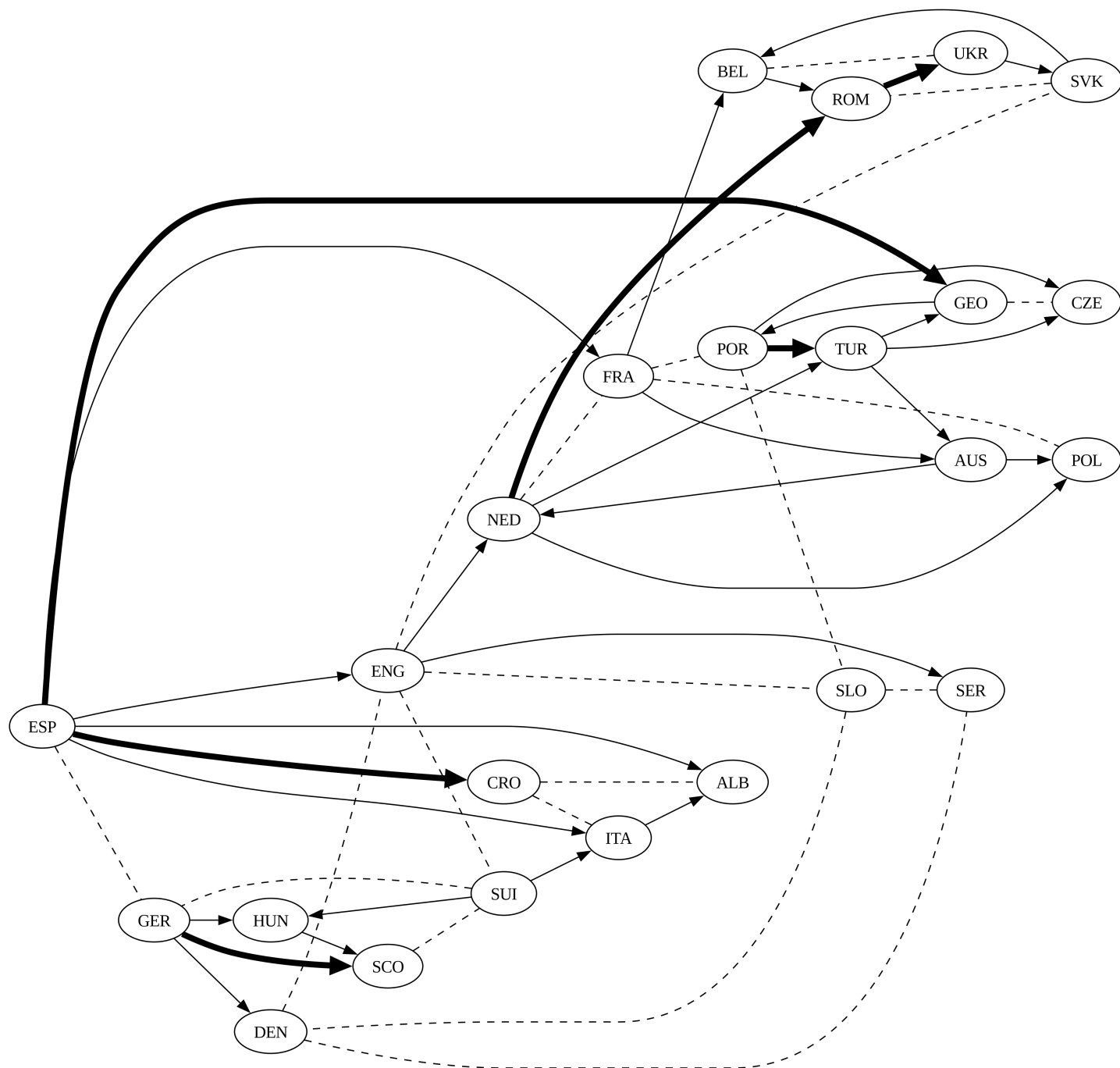
Itt már minden csapat kiértékelhetővé vált. Minden csapathoz tartozik becsült erősség. A becsült erősségek azt mutatják, hogy holtverseny nem alakult ki, ami szintén pozitív tulajdonsága a páros összehasonlítási módszereknek. A becsült erősségeket megvizsgálva láthatjuk, hogy a spanyolok messze kiemelkednek a mezőnyből, és Németország most is a spanyolok után a második helyre került.

4.11.	Csapat	Helyezés	$\hat{\pi}_i$	$\hat{m}_i$
1	ESP	1	0,6471	7,18286
2	ENG	3	0,04845	4,59089
3	NED	5	0,02904	4,07916
4	FRA	4	0,03814	4,35154
5	GER	2	0,09396	5,25324
6	SUI	6	0,02293	3,84294
7	TUR	12	0,00996	3,00915
8	POR	7	0,02258	3,82744
9	AUS	8	0,01498	3,41675
10	BEL	22	0,0008	0,49135
11	SLO	10	0,01116	3,12301
12	SVK	21	0,00086	0,55608
13	ROM	24	0,00049	0
14	ITA	11	0,01068	3,07843
15	GEO	18	0,00296	1,79442
16	DEN	9	0,01388	3,3406
17	UKR	23	0,00051	0,03419
18	HUN	14	0,00546	2,40846
19	SER	13	0,0089	2,89645
20	CRO	17	0,0044	2,19103
21	ALB	16	0,0049	2,29961
22	CZE	19	0,00158	1,16525
23	POL	15	0,00509	2,3381
24	SCO	20	0,0012	0,8894

4.18. táblázat. Az 5 kategóriás modellel az A eset szerint kapott eredmények.

Térjünk át a B esetre.

A 4.8. ábra mutatja az összes mérkőzés eredménye alapján felrajzolt gráfot a B eset alapján. Ez a gráf nem sokban tér el az A eset szerinti gráftól, de épp úgy, mint a 3 kategóriás modellben, itt is igaz, hogy a spanyol és a német csapat más-más éllel van összekötve az A és a B esetben.



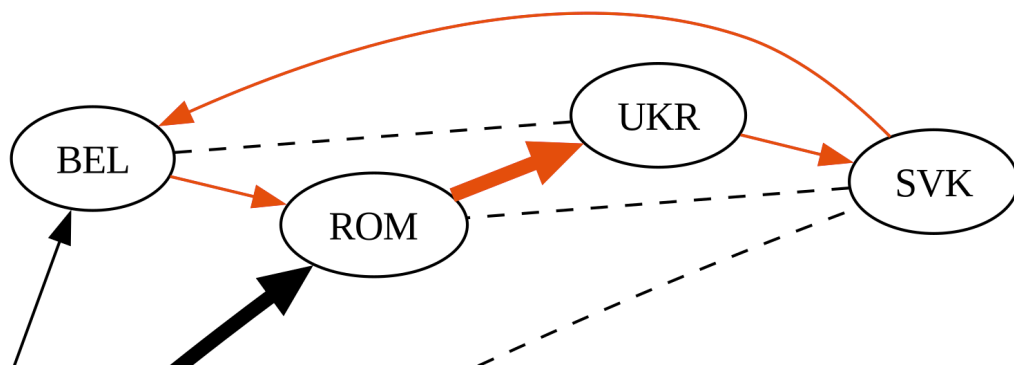
4.8. ábra. A B eset szerint kapott eredmények alapján felrajzolt gráf az 5 kategóriás modellben. ($\Rightarrow$ nagy győzelem, $\rightarrow$ kis győzelem, - - - döntetlen)

Itt is ránézésre látható, hogy a gráf összefüggő. A gráf vizsgálata során,

melyet a 3.2. alfejezetben definiált feltételrendszerrel tehetünk meg, a következőket állapíthatjuk meg:

Az SC 1. feltétel teljesül, mivel létezik nagy vereség (spanyol-horvát mérkőzés), kis vereség (spanyol-angol mérkőzés), illetve döntetlen (horvát-albán mérkőzés) is.

Az SC 2. feltétel teljesül, mivel létezik egy olyan kör, amely csak győzelmekből áll, és az egyik egy nagy győzelem, lásd 4.9.



4.9. ábra. A csupa győzelmekből álló kör a B eset szerinti eredmények az 5 kategóriás modellben.

Az SC 3. feltétel teljesül, mivel minden partíció esetében van vagy köztes döntés (döntetlen vagy kis győzelem) a halmazok között, vagy oda-vissza nagy győzelem/vereség páros. Ezt számítógépes programmal ellenőriztük.

A B eset szerint is kiértékelhetők tehát a meccsek eredményei. A kiértékelés eredményeit a 4.19. táblázatban találhatjuk.

4.11.	Csapat	Helyezés	$\hat{\pi}_i$	$\hat{m}_i$
1	ESP	1	0,38284	6,36071
2	ENG	4	0,035	3,96836
3	NED	9	0,01685	3,23713
4	FRA	8	0,01686	3,23831
5	GER	2	0,33853	6,2377
6	SUI	3	0,08275	4,82893
7	TUR	14	0,00634	2,25942
8	POR	10	0,01662	3,22383
9	AUS	13	0,00696	2,35263
10	BEL	21	0,00106	0,4704
11	SLO	7	0,01776	3,29021
12	SVK	18	0,0025	1,32876
13	ROM	24	0,00066	0
14	ITA	11	0,01225	2,91865
15	GEO	20	0,0018	0,99998
16	DEN	6	0,01839	3,32469
17	UKR	23	0,00084	0,23756
18	HUN	5	0,01847	3,32925
19	SER	12	0,00931	2,64421
20	CRO	17	0,00363	1,70264
21	ALB	15	0,00388	1,76925
22	CZE	22	0,0009	0,3064
23	POL	19	0,00198	1,09542
24	SCO	16	0,00382	1,75384

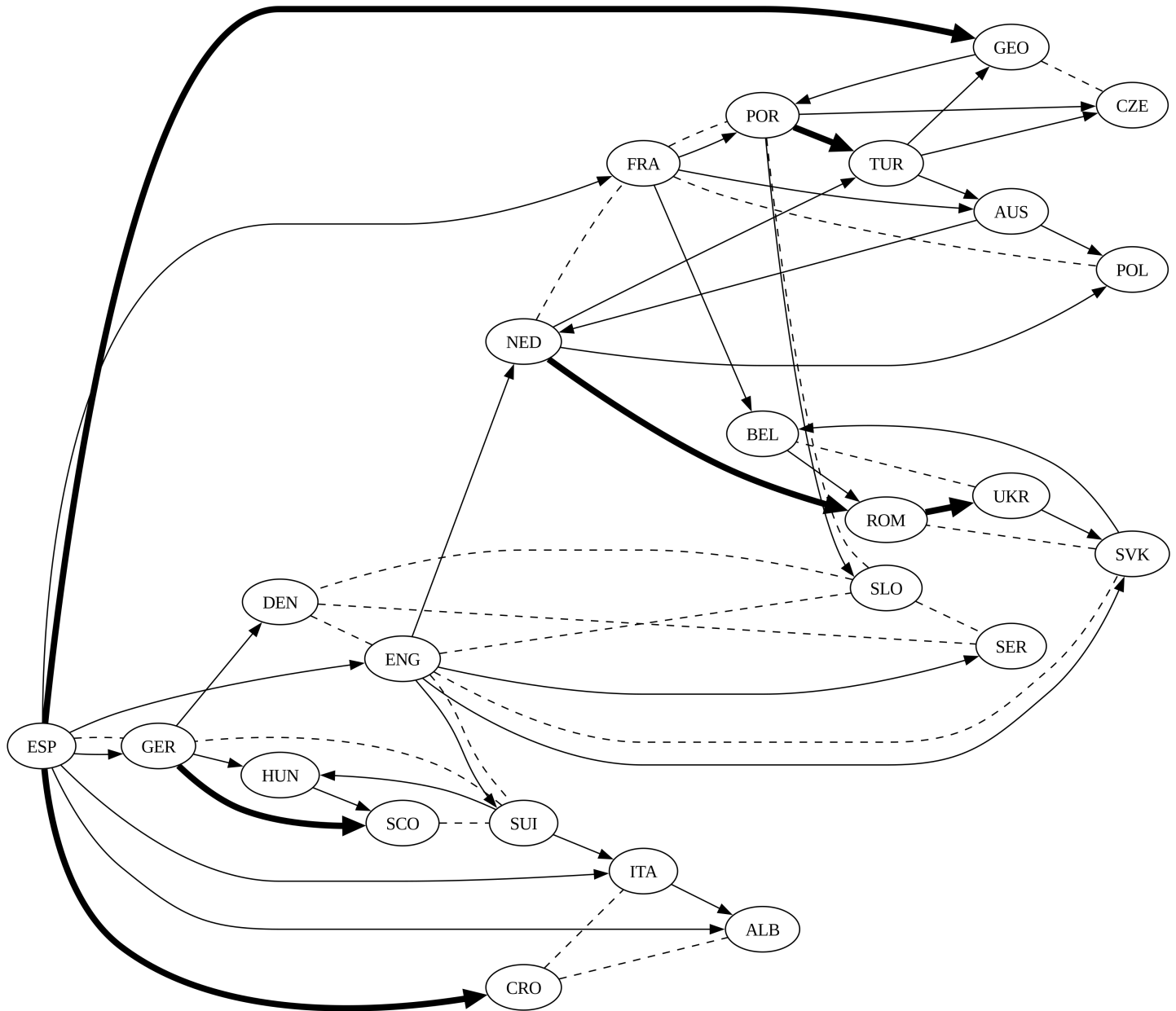
4.19. táblázat. Az 5 kategóriás modellel a B eset szerint elért eredmények.

Az első helyen Spanyolország, a második helyen Németország tanyázik. A svájci csapat itt a harmadik helyre került, mivel a második helyezett német csapattal játszottak egy döntetlent, a németek pedig az első helyen szereplő spanyolokkal szemben jutottak szintén döntetlen eredményre. A magyar csapat nagyon sokat lépett előre, mivel a svájciakkal döntetlent játszó skótokat sikerült legyőzniük egy gólkülönbséggel. A holland csapat viasszaszorult a ki-

lencedik helyre, mivel egy, az A esethez képest gyengébb angol csapat verte meg őket, és csak a török, illetve a sereghajtó román csapatot tudták legyőzni, utóbbit nagy gólkülönbséggel. A belga csapat itt is gyengének tekinthető, de mivel a szlovák csapat itt erősebbnek számít az angolok elleni döntetlen miatt, így ők egyet előreléptek, a huszonkettedikről a huszonegyedik helyre.

Térjünk át a C esetre, mely egyben meghatározza a D és E esetek kiértékelhetőségét.

A C esethez tartozó gráf eredményeit a 4.10. ábra mutatja. Ez a gráf egyben a D és az E esetbeli gráfot is jelenti. Ez a gráf az előző gráfok összeolvasztásával kapható meg. Ahol az A és a B esetbeli gráfok közt különbség van, ott a 4.10 gráfban döntetlen és győzelem él is megtalálható.



4.10. ábra. A C, D és E esetek szerint kapott eredmények alapján felrajzolt gráf az 5 kategóriás modellben. ($\Rightarrow$ nagy győzelem, $\rightarrow$ kis győzelem, - - - döntetlen)

Belátható továbbá, hogy a 3.2 alfejezetben található SC 1., SC 2. és SC 3. feltételek is mind teljesülnek. Ebből következtetve a C, D és E eset szerint is kiértékelhetők a meccsek eredményei. A kiértékelés eredményeit a 4.20., a 4.21. és a 4.22. táblázatban találhatjuk.

4.11.	Csapat	Helyezés	$\hat{\pi}_i$	$\hat{m}_i$
1	ESP	1	0,41094	5,78587
2	ENG	4	0,04764	3,6311
3	NED	7	0,02746	3,08027
4	FRA	6	0,03116	3,20666
5	GER	2	0,23303	5,21861
6	SUI	3	0,06964	4,0108
7	TUR	14	0,00998	2,06763
8	POR	8	0,02273	2,89095
9	AUS	12	0,01248	2,2918
10	BEL	21	0,00199	0,45786
11	SLO	10	0,01702	2,60161
12	SVK	17	0,00412	1,18304
13	ROM	24	0,00126	0
14	ITA	5	0,03549	3,33656
15	GEO	20	0,00285	0,81483
16	DEN	9	0,01948	2,73661
17	UKR	22	0,00153	0,19588
18	HUN	11	0,01578	2,52622
19	SER	13	0,01081	2,14755
20	CRO	16	0,00769	1,80738
21	ALB	15	0,00811	1,86083
22	CZE	23	0,0015	0,17389
23	POL	18	0,00382	1,1073
24	SCO	19	0,0035	1,019

4.20. táblázat. Az 5 kategóriás modellel a C eset szerint elért eredmények.

Mindhárom (C, D, E) eset szerint az első a spanyol, a második a német, a harmadik pedig a svájci csapat, az angolok pedig a negyedik helyre kerültek. A magyar csapat az előkelő tizedik, illetve tizenegyedik helyen van, köszönhetően annak, hogy kis vereséget szenvedett a második helyen lévő német csapattól. A 4.20, a 4.21 és 4.22 táblázatok által mutatott D és E esetben a szlovén csapat hatodik és nyolcadik és tizedik helyezett attól függően, hogy a hosszabbításban

elért eredményét hogyan értelmezzük. Ennek azért van jelentősége, mert csak a portugáloktól kaptak ki 11-es rúgások folytán.

4.11.	Csapat	Helyezés	$\hat{\pi}_i$	$\hat{m}_i$
1	ESP	1	0,48483	6,54067
2	ENG	4	0,03832	4,00273
3	NED	6	0,01946	3,32509
4	FRA	5	0,02019	3,36191
5	GER	2	0,24741	5,86793
6	SUI	3	0,06453	4,52399
7	TUR	14	0,00722	2,33417
8	POR	7	0,01848	3,27336
9	AUS	13	0,00809	2,44806
10	BEL	21	0,00115	0,49275
11	SLO	8	0,01742	3,21421
12	SVK	18	0,00261	1,31667
13	ROM	24	0,0007	0
14	ITA	11	0,01254	2,88577
15	GEO	20	0,00202	1,06222
16	DEN	9	0,01727	3,20561
17	UKR	23	0,0009	0,25241
18	HUN	10	0,01349	2,95862
19	SER	12	0,00924	2,57989
20	CRO	16	0,00394	1,72795
21	ALB	15	0,00423	1,79792
22	CZE	22	0,00099	0,34617
23	POL	19	0,00228	1,17918
24	SCO	17	0,00272	1,35822

4.21. táblázat. Az 5 kategóriás modellel a D eset szerint elért eredmények.

A 4.21. táblázat által mutatott D esetben szembeűnő az olasz csapat tizenegyedik helyezése, mely betudható annak, hogy csak az albánokat sikerűlt legyőzniűk, kis gólkűlönbséggel, illetve attól a svájci csapattól kaptak ki, akik

csak a 11-es rúgások során kaptak ki a negyedik helyezett angoloktól.

4.11.	Csapat	Helyezés	$\hat{\pi}_i$	$\hat{m}_i$
1	ESP	1	0,4601	6,85872
2	ENG	4	0,03885	4,38696
3	NED	9	0,01688	3,55329
4	FRA	8	0,0171	3,56642
5	GER	2	0,26797	6,31816
6	SUI	3	0,07353	5,02497
7	TUR	14	0,00649	2,59791
8	POR	7	0,01715	3,56929
9	AUS	13	0,0072	2,70083
10	BEL	22	0,00078	0,48299
11	SLO	6	0,01798	3,61641
12	SVK	20	0,00143	1,08411
13	ROM	24	0,00048	0
14	ITA	11	0,01303	3,29474
15	GEO	19	0,00193	1,38465
16	DEN	5	0,01833	3,63577
17	UKR	23	0,00059	0,19183
18	HUN	10	0,01553	3,47028
19	SER	12	0,00991	3,02048
20	CRO	16	0,00409	2,13461
21	ALB	15	0,00439	2,20714
22	CZE	21	0,00096	0,68425
23	POL	18	0,00209	1,46596
24	SCO	17	0,00321	1,89313

4.22. táblázat. Az 5 kategóriás modellel az E eset szerint elért eredmények.

A 4.22. táblázat által mutatott E esetben a szlovén csapat méginkább előretört, így a hatodik helyet érték el. Ez annak köszönhető, hogy a negyedik helyezett angol csapattal játszott döntetlenük ebben az esetben nagyobb súllyal rendelkezik, mint az előzőben. A dán csapat az ötödik helyezett lett, mivel

az angolokkal játszottak egy döntetlent, továbbá a német csapattól is csak kis gólkülönbséggel kaptak ki. A magyar csapat itt is a tizedik helyezett lett, mivel a svájciakkal döntetlent játszó skótokat megverték. Meglepő látvány továbbá a kilencedik helyezett holland csapat, akiket a tizenharmadik helyezett osztrák csapat győzött le 1 gólkülönbséggel, majd a negyedik helyezett angolok ejtették ki őket az elődöntőkben, szintén kis gólkülönbséggel.

Annak érdekében, hogy még könnyebben össze tudjuk hasonlítani a kiértékelési eredményeket, egy közös 4.23. táblázatban is megjelenítettük a sorrendeket.

Összefoglalva: az 5 kategóriás kiértékelések eredményei alapján a legerősebb csapat a spanyol csapat, őt követi a német, majd az angol vagy a svájci, attól függően, hogy a hosszabbításban eldőlt meccsek eredményét hogyan értelmezzük. Egyöntetűen a leggyengébb csapat a román csapat. Nem sokkal jobb az ukrán, a belga, a cseh, a grúz/georgiai és a szlovák csapat.

A magyar csapat a középmezőnyben helyezkedik az A, C, D és E esetek szerinti eredmények alapján, viszont amikor a hosszabbításban eldőlt mérkőzések eredményét döntetlennek tekintjük, akkor nagyon előkelő helyre kerül. Ennek oka, hogy a második legerősebb, a német csapat ellen kis vereséget szenvedtünk, a németek pedig az első helyezett spanyolokkal szemben döntetlent játszottak.

Ha javaslatot kellene tenni arra, hogy melyik értelmezést használjuk a kiértékeléskor, akkor az E esetet javasolnám, mert ekkor minden lejátszott mérkőzés ugyanakkora súllyal szerepel a likelihood függvényben.

4.11.	Csapat	3 kategória			5 kategória				
		A	B	C	A	B	C	D	E
1	ESP	—	1	1	1	1	1	1	1
2	ENG	2	6	4	3	4	4	4	4
3	NED	4	4	6	5	9	7	6	9
4	FRA	3	3	5	4	8	6	5	8
5	GER	1	2	2	2	2	2	2	2
6	SUI	9	7	3	6	3	3	3	3
7	TUR	6	8	8	12	14	14	14	14
8	POR	5	5	7	7	10	8	7	7
9	AUS	7	9	11	8	13	12	13	13
10	BEL	—	18	20	22	21	21	21	22
11	SLO	10	10	10	10	7	10	8	6
12	SVK	—	12	15	21	18	17	18	20
13	ROM	—	19	19	24	24	24	24	24
14	ITA	—	—	—	11	11	5	11	11
15	GEO	12	14	16	18	20	20	20	19
16	DEN	8	11	9	9	6	9	9	5
17	UKR	—	17	17	23	23	22	23	23
18	HUN	14	15	12	14	5	11	10	10
19	SER	11	13	13	13	12	13	12	12
20	CRO	—	—	—	17	17	16	16	16
21	ALB	—	—	—	16	15	15	15	15
22	CZE	15	21	21	19	22	23	22	21
23	POL	13	16	18	15	19	18	19	18
24	SCO	16	20	14	20	16	19	17	17

4.23. táblázat. A 3 és 5 kategóriás kiértékelések eredményei.

4.4.2. A különböző kiértékelési módszerek eredményeinek korrelációja

Az előző, 4.3.2. alfejezetben ismertetett súly- és rangkorrelációk kiszámításával fogjuk elemezni az 5 kategóriás kiértékelések eredményeinek összefüggősségeit

egymással, illetve a 4.11. táblázat rangsorával. Az eredményeket a 4.24., 4.25. és 4.26 táblázatban láthatjuk.

	A	B	C	D	E
A	1	0,81233	0,92129	0,93878	0,91554
B	0,81233	1	0,97228	0,96335	0,97794
C	0,92129	0,97228	1	0,99702	0,99785
D	0,93878	0,96335	0,99702	1	0,99803
E	0,91554	0,97794	0,99785	0,99803	1

4.24. táblázat. Az 5 kategóriás modellben az A, B, C, D és E esetek erősségeinek összehasonlítása, Pearson-féle súlykorrelációt használva.

Ahogy várható, az A és a C, illetve a B és a C eset között erős korreláció tapasztalható mindhárom korrelációs együttható esetén. Ennek az az oka, hogy a mérkőzések eredményeinek értelmezése erősen hasonlít, míg az A és a B eset szerinti értelmezések sokkal inkább különböznek. Az A és a B eset korrelációs együtthatója 80% körül van a Pearson-féle esetben, míg a Spearman-féle rangkorreláció nagyobb mint a Kendall-féle rangkorreláció. A Kendall-féle rangkorreláció az A és a B eset között 0,76, a Spearman-féle korreláció pedig 0,89. Az előbbi inkább számít közepes, mint erős korrelációnak. A 4.24. táblázatban azért nem szerepelnek a nemhivatalos táblázattal való összevetések, mert a táblázat nem tartalmaz erősségeket. A rangsorok korrelációi a nemhivatalos és a páros összehasonlítási módszerek között a Kendall-féle esetben már 0,5 körüli, illetve 0,5 alatti értékek, ami gyenge korrelációnak értelmezhető. A Spearman-féle esetben a korrelációs értékek a 0,58-0,71 tartományba tartoznak. A C, D és E esetek nagyon erősen összefüggenek 0,997 feletti értékkel. Ez azt mutatja, hogy amennyiben a hosszabbításban eldőlt mérkőzéseket különböző súlyokkal ugyan, de mind győzelemként, mind döntetlenként figyelembe vesszük, akkor a kiértékelések eredményei nagyon hasonlóak lesznek.

	A	B	C	D	E	4.11.
A	1	0,76087	0,81884	0,84058	0,81159	0,55072
B	0,76087	1	0,84058	0,9058	0,92029	0,41304
C	0,81884	0,84058	1	0,92029	0,84783	0,51449
D	0,84058	0,9058	0,92029	1	0,91304	0,50725
E	0,81159	0,92029	0,84783	0,91304	1	0,42029
4.11.	0,55072	0,41304	0,51449	0,50725	0,42029	1

4.25. táblázat. Az 5 kategóriás modellben az A, B, C, D és E esetek rangjainak összehasonlítása egymással és a 4.11. nemhivatalos ponttáblázat rangsorával, Kendall-féle rangkorrelációt használva.

	A	B	C	D	E	4.11.
A	1	0,89739	0,94174	0,9513	0,93652	0,71043
B	0,89739	1	0,9487	0,97217	0,98	0,58696
C	0,94174	0,9487	1	0,97652	0,95652	0,68
D	0,9513	0,97217	0,97652	1	0,98	0,68087
E	0,93652	0,98	0,95652	0,98	1	0,60174
4.11.	0,71043	0,58696	0,68	0,68087	0,60174	1

4.26. táblázat. Az 5 kategóriás modellben az A, B, C, D és E esetek rangjainak összehasonlítása egymással és a 4.11. nemhivatalos ponttáblázat rangsorával, Spearman-féle rangkorrelációt használva.

A 4.25. és a 4.26. táblázatban található rangkorrelációk alapján látható, hogy a 4.11. ponttáblázat rangsora közepes összefüggést mutat a páros összehasonlítások segítségével kapott rangsorokkal. Az A eset áll legközelebb a ponttáblázat rangsorához. Az A eset és a B eset szintén gyengébben korrelál egymással, mint a C, D és E esetekkel, hasonlóan a 3 opciós modellben kapott eredményekhez.

	A eset 3 és 5 kategória
Pearson-féle súlykorreláció	0,88482
Kendall-féle rangkorreláció	0,83333
Spearman-féle rangkorreláció	0,93824

4.27. táblázat. A 3 és 5 kategóriás eredmények összehasonlítása az A eset adatai szerint.

	B eset 3 és 5 kategória
Pearson-féle súlykorreláció	0,80879
Kendall-féle rangkorreláció	0,6
Spearman-féle rangkorreláció	0,78052

4.28. táblázat. A 3 és 5 kategóriás eredmények összehasonlítása a B eset adatai szerint.

	C eset 3 és 5 kategória
Pearson-féle súlykorreláció	0,96005
Kendall-féle rangkorreláció	0,87619
Spearman-féle rangkorreláció	0,96104

4.29. táblázat. A 3 és 5 kategóriás eredmények összehasonlítása a C eset adatai szerint.

A 4.27., 4.28. és 4.29. táblázatokban a 3 kategóriában kiértékelhető rész-halmaz, illetve a kiértékelhető rész-halmaz méretére lezsugorított 5 kategóriás kiértékelés eredménye van összehasonlítva a három, előzőleg is alkalmazott korreláció segítségével. Ekkor a 3 kategóriás A esetet az 5 kategóriás A esethez hasonlítottuk, és ugyanígy tettünk a B és a C eseteknél is. Azért végeztük el ezt a vizsgálatot, hogy lássuk, mennyit módosít az eredményeken a több információ beépítése azon csapatoknál, amelyek a 3 kategóriás modellben is kiértékelhetők voltak. Láthatjuk, hogy a C esetben magasak a korrelációs együtthatók, az A és a B esetben kicsit alacsonyabbak. Ez is azt támasztja alá, hogy az adatok értelmezését a C (vagy D vagy E) eset szerint érdemes végezni. Ez azt jelenti, hogy a hosszabbítás során eldőlt mérkőzéseket érdemes mind döntetlenként, mind győzelemként figyelembe venni.

5. fejezet

Informatikai megvalósítás

A dolgozatban látható kiértékelések eredményeinek kiszámítását, továbbá a különböző esetek kiértékelhetőségének vizsgálatát számítógépes programmal végeztem. A programot a C# programozási nyelvben írtam, mivel könnyű használni, illetve a gyorsasága is kielégítő.

A csapatok neveit, illetve a mérkőzések eredményeit szövegfájlokban tároltam, amelyeket a program a futás alatt olvas be. A program tartalmazza a fixpont iterációs eljárást, a loglikelihood függvényértékek és parciális deriváltak számolását, a kiértékelhetőség vizsgálat részeként az összefüggőség vizsgálatot, a körkeresést és az irányított gráf erős összefüggőségének vizsgálatát. A program képes gyorsan kiértékelni páros összehasonlítási modellekből származó adatokat. A program továbbá tud generálni L^AT_EX-táblázatokat a kiértékelések eredményéből.

A 3.2. alfejezetben található kiértékelhetőség vizsgálat 5 kategóriás modell esetén az alábbi módon van implementálva a programkódban:

```

1  ...
2  public void KiertekeleshetosegEllenorzes(){
3      OsszefuggosegEllenorzes();
4      if(osszefuggo==false){
5          return;
6      }
7      bool nagyVereseg=false;
8      bool kisVereseg=false;
9      bool dontetlen=false;
10     for(int i=0;i<Elemszam-1;i++){
11         for(int j=i+1;j<Elemszam;j++){
12             int u = Hasonlit(i,j);

```

```

13         if(A[u,1-1]>0 || A[u,5-1]>0){
14             nagyVereseg=true;
15         }
16         if(A[u,2-1]>0 || A[u,4-1]>0){
17             kisVereseg=true;
18         }
19         if(A[u,3-1]>0){
20             dontetlen=true;
21         }
22         if(nagyVereseg && kisVereseg && dontetlen){
23             goto ElsoFeltetelBreak;
24         }
25     }
26 }
27 if(!(dontetlen && nagyVereseg && kisVereseg)){
28     kiertekekelheto=false;
29     kiertVizsgalva=true;
30     return;
31 }
32 ElsoFeltetelBreak:
33 bool[] meglatogatott = new bool[ELEMSZAM];
34 Stack<int> ut = new Stack<int>();
35 bool[] veremTagok = new bool[ELEMSZAM];
36 for(int i=1;i<ELEMSZAM;i++){
37     if(meglatogetott[i]==false){
38         KorKereses(i, meglatogatott, ut, veremTagok);
39     }
40 }
41 if(vanKor==false){
42     kiertekekelheto=false;
43     kiertVizsgalva=true;
44     return;
45 }
46 bool erosenOsszefuggo = ErosOsszefuggesVizsgalat();
47 if(erosenOsszefuggo==false){
48     kiertekekelheto=false;
49     kiertVizsgalva=true;
50     return;
51 }
52 kiertekekelheto=true;
53 }
54 ...

```

A körkeresésre egy mélységi bejárás alapú algoritmust használtam, amely a megtalált köröket ellenőrzi, és amennyiben talál egy olyan kört, amely megfelel az SC 2. feltételnek, abbahagyja a keresést. Az implementációja így néz ki:

```

1 ...
2 public readonly bool TalaltKorEllenorzes(List<int> kor){
3     int kicsi = 0;
4     int nagy = 0;

```

```

5     for(int i=0;i<kor.Count;i++){
6         int u = kor[i];
7         int v = kor[(i + 1) % kor.Count];
8         if(kisGyozelem[u, v]){
9             kicsi++;
10        }
11        if(nagyGyozelem[u, v]){
12            nagy++;
13        }
14    }
15    return nagy>=1;
16 }
17 public void KorKereses(int u, bool[] meglatogatott, Stack<int> utvonal, bool[]
    veremTagok){
18     if(vanKor==true){
19         return;
20     }
21     meglatogatott[u] = true;
22     veremTagok[u] = true;
23     utvonal.Push(u);
24     for(int v=1;v<Elemszam;v++){
25         if(kisGyozelem[u, v]==true || nagyGyozelem[u, v]==true){
26             if (!meglatogetott[v]){
27                 KorKereses(v, meglatogatott, utvonal, veremTagok);
28                 if(vanKor==true){
29                     return;
30                 }
31             }
32             else if(veremTagok[v]==true){
33                 List<int> kor = new List<int>();
34                 foreach(int csucs in utvonal){
35                     kor.Add(csucs);
36                     if (csucs == v){
37                         break;
38                     }
39                 }
40                 kor.Reverse();
41                 if(TalaltKorEllenorzes(kor)==true){
42                     vanKor=true;
43                     return;
44                 }
45             }
46         }
47     }
48     utvonal.Pop();
49     veremTagok[u] = false;
50 }
51 ...

```

A 3.1. és 3.2. alfejezetekben található, módosított partícióellenőrzés helyett gráf erős összefüggőséget vizsgáló SC 3. feltételt az 5 kategóriás modell

kiértékelhetőség vizsgálatában implementáltam. Az implementációra a Kosajaru algoritmus módosított változatát használtam. A hagyományos Kosajaru algoritmus meghatározza, hogy egy irányított gráf hány erősen összefüggő komponensből áll. Az általam alkalmazott egyszerűsítése meghatározza, hogy egy erősen összefüggő komponensből áll, vagy sem. Amennyiben ez a feltétel teljesül, akkor az irányított gráf erősen összefüggő, amennyiben viszont nem teljesül, akkor nem erősen összefüggő. A programkódban való implementációja a következőképpen néz ki:

```

1  ...
2  public void DFS(int v, bool[] visited){
3      visited[v] = true;
4      for(int i=0;i<ELEMSZAM;i++){
5          if(grafSC[v,i] && !visited[i]){
6              DFS(i, visited);
7          }
8      }
9  }
10 public void DFSTranszponalt(int v, bool[] meglatogatott, bool[,] transzponalt)
    {
11     meglatogatott[v] = true;
12     for(int i=0;i<ELEMSZAM;i++){
13         if(transzponalt[v, i] && !meglatogetott[i]){
14             DFSTranszponalt(i, meglatogatott, transzponalt);
15         }
16     }
17 }
18 public bool[,] Transzponalt()
19 {
20     bool[,] transzponalt = new bool[ELEMSZAM, ELEMSZAM];
21     for(int i=0;i<ELEMSZAM;i++){
22         for(int j=0;j<ELEMSZAM;j++){
23             if(grafSC[i, j]){
24                 transzponalt[j, i] = true;
25             }
26         }
27     }
28     return transzponalt;
29 }
30 public bool ErosOsszefuggesVizsgalat(){
31     bool[] meglatogatott = new bool[ELEMSZAM];
32     DFS(0, meglatogatott);
33     for(int i=0;i<ELEMSZAM;i++){
34         if (!meglatogetott[i])
35             return false;
36     }

```

```

37     bool[,] transzponaltMatrix = Transzponalt();
38     for(int i=0;i<ELEMSZAM;i++){
39         meglatogatott[i] = false;
40     }
41     DFSTranszponalt(0, meglatogatott, transzponaltMatrix);
42     for(int i=0;i<ELEMSZAM;i++){
43         if(!meglatogetott[i]){
44             return false;
45         }
46     }
47     return true;
48 }
49 ...

```

A fixpont iterációs eljárás 5 kategóriás modell esetén a következőképpen néz ki:

```

1  ...
2  public void FixpontIteracio(){
3      theta=2;
4      beta=3;
5      for(int i=0;i<ELEMSZAM;i++){
6          pi[i]=1.0/ELEMSZAM;
7      }
8      double[] oldPi = new double[ELEMSZAM];
9      double oldTheta;
10     double oldBeta;
11     double eps12=0.0000001;
12     double epsderiv=0.0000001;
13     for(int k=0;k<1000;k++){
14         for(int i=0;i<ELEMSZAM;i++){
15             oldPi[i]=pi[i];
16         }
17         oldTheta=theta;
18         oldBeta=beta;
19         double sum=0;
20         for(int i=0;i<ELEMSZAM;i++){
21             pi[i]=Pi_kFx(i);
22             sum+=pi[i];
23         }
24         theta=ThetaFx();
25         beta=BetaFx();
26         for(int i=0;i<ELEMSZAM;i++){
27             pi[i]=pi[i]/sum;
28         }
29         double l2norma=0;
30         for(int i=0;i<ELEMSZAM;i++){
31             l2norma+=Math.Pow(oldPi[i],2);
32         }
33         l2norma+=Math.Pow(oldTheta,2);
34         l2norma+=Math.Pow(oldBeta,2);
35         l2norma=Math.Sqrt(l2norma);

```

```

36         if(l2norma<eps12){
37             Console.WriteLine(k);
38             break;
39         }
40         bool kisD=true;
41         for(int i=0;i<ELEMSZAM;i++){
42             if(Math.Abs(DPi_k(i))>epsderiv){
43                 kisD=false;
44             }
45         }
46         if(Math.Abs(DTheta())>epsderiv){
47             kisD=false;
48         }
49         if(Math.Abs(DBeta())>epsderiv){
50             kisD=false;
51         }
52         if(kisD==true){
53             Console.WriteLine(k);
54             break;
55         }
56     }
57 }
58 ...

```

A 3 kategóriás modell implementációja a módosíthatlan SC 3. feltétellel, a C eset szerint kiértékelhető 21 csapatot magában foglaló részhalmaz adatainak kiértékelését összesen 0,620 másodperc alatt képes megtenni. A 2.1.1 alfejezetben lévő $\varepsilon_1, \varepsilon_2$ küszöbértékeket 10^{-7} , az $1/\varepsilon_3$ iterációs szám határt pedig 1000-nek választottam. Összesen 267 iteráció kellett a maximumhely megtalálásához a 3 kategóriás modellben. Az 5 kategóriás modell a C eset szerinti adatokat, 24 csapattal, összesen 0,062 másodpercen belül értékelte ki a program. Ekkor azonos határértékeket alkalmazva 567 iterációs lépésre volt szükség. Ebből látható, hogy a módosított SC 3. feltétel drasztikusan tudja gyorsítani a kiértékelhetőség vizsgálatát. A kiértékelések egy AMD Ryzen 3 5425U processzorral rendelkező laptopon lettek végrehajtva.

A megvalósított gyors futású program alkalmas lehet további, szimulációk futtatásával végezhető strukturális vizsgálatokra.

6. fejezet

Összefoglalás

A labdarúgásban a hagyományosan használt ponttáblázatok bár könnyen értelmezhetők, de erős limitációik vannak. A pontozás nem veszi figyelembe a csapatok erősségét, továbbá Európa-bajnokságokon, világbajnokságokon és hasonló tornákon csak a csoportkörökben tartanak nyilván ilyenemű sorrendeket. Az UEFA honlapján nem is található a vélt erősségeket bemutató lista például a labdaúgó Európa-bajnoksággal kapcsolatban sem, pedig az egyes csapatok erősségének megbecslése sokak számára érdekes információt biztosítana. Ezt a feladatot végeztem el a Tudományos Diákköri Konferenciára készült munkámban.

Sporteredmények kiértékelésére használnak 3 opciós Bradley-Terry modellt, ez azonban a 2024-es labdarúgó Európa-bajnokság mérkőzéseinek kiértékelésére nem bizonyult alkalmasnak. Ezért a munkám során implementált 3 opciós Bradley-Terry modellt továbbfejlesztettem 5 opciós modellé. A 3 opciós modell esetén ellenőriztem a kiértékelhetőség feltételeit, a létrehozott fixpont iterációs módszert a kiértékelhető részhalmaz elemzésére felhasználtam. Az 5 opciós modellre kidolgoztam egy numerikus módszert az optimalizálásra, ellenőriztem a kiértékelhetőség feltételeit, a hosszabbításban eldölt mérkőzések eredményeit többféleképpen értelmeztem, és azt tapasztaltam, hogy jelentős különbség alakul ki, ha azokat győzelemként vagy döntetlenként értelmezzük. Ha viszont győzelemként és döntetlenként is figyelembe vesszük őket különböző súlyozásokkal, akkor a súlyozások megválasztása a kiértékelési eredményeket már alig befolyásolja. Mivel az E esetben az összes lejátszott mérkőzés jelentősége ugyanakkora a számolások során, ezért az E esetet javaslom az eredmények értelmezésére.

A munkám során elvégzett elemzések azt támasztják alá, hogy a labdarúgó Európa-bajnokságot megnyerő spanyol csapat erőssége messze meghaladja a többi ország erősségét. Mindegyik értelmezés esetén második Németország, annak ellenére, hogy nem sikerült a döntőbe jutnia. A módszer kimondottan pozitív tulajdonsága, hogy sikerült azonosítania a második legerősebb csapatot, még akkor is, ha a legerősebb csapattal szembeni veresége miatt korai kiesés lett a sorsa. A leggyengébb csapatokat is egyértelműen meg lehet határozni. A módszer jól értelmezhető sorrendet ad a csoportkörben elvérző csapatok esetében is, elkerülve a holtversenyt. A középmezőnyben találhatjuk a legnagyobb változékonyságot.

A kidolgozott módszer kifinomult elemzést tesz lehetővé sportbajnokságok kiértékelhetőségében, és alkalmazható lenne a Nemzetek Ligája vagy az Európa Liga esetén is. A TDK munka készítése során végzett kutatás alapján bátran ajánlom alkalmazását.

Irodalomjegyzék

- [1] Cheng, K. E., McHugh, J. A., and Deek, F. P. (2013). On the use of paired comparisons to construct group preference scales for decision making. *Group Decision and Negotiation*, 22, 519-540.
- [2] YDorado, R., Gómez-Moreno, A., Torres-Jiménez, E., and López-Alba, E. (2014). An AHP application to select software for engineering education. *Computer Applications in Engineering Education*, 22(2), 200-208.
- [3] Stein, A.; Aryal, J.; Gort, G. Use of the Bradley-Terry model to quantify association in remotely sensed images. *IEEE Trans. Geosci. Remote Sens.* 2005, 43, 852–856. <https://doi.org/10.1109/TGRS.2005.843569>
- [4] Badri, M. A., and Abdulla, M. H. (2004). Awards of excellence in institutions of higher education: an AHP approach. *International Journal of Educational Management*, 18(4), 224-242.
- [5] Jeon, J. J., and Kim, Y. (2013). Revisiting the Bradley-Terry model and its application to information retrieval. *Journal of the Korean Data and Information Science Society*, 24(5), 1089-1099. <http://dx.doi.org/10.7465/jkdi.2013.24.5.1089>
- [6] Tutz, G., and Schauburger, G. (2015). Extended ordered paired comparison models with application to football data from German Bundesliga. *AStA Advances in Statistical Analysis*, 99, 209-227. <https://doi.org/10.1007/s10182-014-0237-1>
- [7] Thurstone, L.L. (1927) A law of comparative judgment. *Psychol. Rev.* 34, 273–286. <https://doi.org/10.1037/h0070288>.

- [8] Bradley, R. A., and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, 39(3/4), 324-345.
- [9] Rao, P. V., and Kupper, L. L. (1967). Ties in paired-comparison experiments: A generalization of the Bradley-Terry model. *Journal of the American Statistical Association*, 62(317), 194-204. <https://doi.org/10.1080/01621459.1967.10482901>.
- [10] Glenn, W. A., and David, H. A. (1960). Ties in paired-comparison experiments using a modified Thurstone-Mosteller model. *Biometrics*, 16(1), 86-109.
- [11] Davidson, R. R. (1970). On extending the Bradley-Terry model to accommodate ties in paired comparison experiments. *Journal of the American Statistical Association*, 65(329), 317-328.
- [12] Agresti, A. (1992). Analysis of ordinal paired comparison data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 41(2), 287-297. <https://doi.org/10.2307/2347562>.
- [13] Orbán-Mihálykó, É., Mihálykó, C.,; Koltay, L. (2019). Incomplete paired comparisons in case of multiple choice and general log-concave probability density functions. *Central European Journal of Operations Research*, 27, 515-532., <https://doi.org/10.1007/s10100-018-0568-1>
- [14] Ford Jr, L. R. (1957). Solution of a ranking problem from binary comparisons. *The American Mathematical Monthly*, 64(8P2), 28-33. <https://doi.org/10.1080/00029890.1957.11989117>.
- [15] Hunter, D. R. (2004). MM algorithms for generalized Bradley-Terry models. *The Annals of Statistics*, 32(1), 384-406.
- [16] Gyarmati, L., Orbán-Mihálykó, É., and Mihálykó, C. (2023). Comparative analysis of the existence and uniqueness conditions of parameter estimation in paired comparison models. *Axioms*, 12(6), 575. <https://doi.org/10.3390/axioms12060575>

- [17] Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15(3), 234–281. [https://doi.org/10.1016/0022-2496\(77\)90033-5](https://doi.org/10.1016/0022-2496(77)90033-5).
- [18] Gyarmati, L., Orbán-Mihálykó, É., Mihálykó, C., and Vathy-Fogarassy, Á. (2023). Aggregated rankings of top leagues’ football teams: Application and comparison of different ranking methods. *Applied Sciences*, 13(7), 4556. <https://doi.org/10.3390/app13074556>
- [19] Temesi, J., Szádóczki, Z., and Bozóki, S. (2024). Incomplete pairwise comparison matrices: Ranking top women tennis players. *Journal of the Operational Research Society*, 75(1), 145-157. <https://doi.org/10.1080/01605682.2023.2180447>.
- [20] Spearing, H., Tawn, J., Irons, D., and Paulden, T. (2023). Modeling intransitivity in pairwise comparisons with application to baseball data. *Journal of Computational and Graphical Statistics*, 32(4), 1383-1392. <https://doi.org/10.1080/10618600.2023.2177299>
- [21] Csató, L. (2017). On the ranking of a Swiss system chess team tournament. *Annals of Operations Research*, 254(1), 17-36.
- [22] Orbán-Mihálykó, É., Mihálykó, C., and Gyarmati, L. (2022). Application of the generalized Thurstone method for evaluations of sports tournaments’ results. *Knowledge*, 2(1), 157-166.
- [23] A 2024. évi UEFA labdarúgó Európa-bajnokság hivatalos eredményei, <https://www.uefa.com/euro2024/fixtures-results/>
- [24] Gyarmati, L., Mihálykó, C., Orbán-Mihálykó, É. (2024). Algorithm for Option Number Selection in Stochastic Paired Comparison Models. *Algorithms*, 17(9), 410. <https://doi.org/10.3390/a17090410>
- [25] Walpole, R. E., Myers, R. H., Myers, S. L., Ye, K. (1993). *Probability and statistics for engineers and scientists*, New York: Macmillan.
- [26] Abdi, H. (2007). The Kendall rank correlation coefficient. *Encyclopaedia of Measurement and Statistics* (pp. 508-510).