

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Rok Fujs

**Ocenjevanje verjetnosti zmage
favorita v različnih oblikah tekmovanj**

DIPLOMSKO DELO
UNIVERZITETNI ŠTUDIJSKI PROGRAM RAČUNALNIŠTVA
IN INFORMATIKA

MENTOR: doc. dr. Erik Štrumbelj

Ljubljana 2014

Rezultati diplomskega dela so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavlanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

Besedilo je oblikovano z urejevalnikom besedil \LaTeX .

Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Tematika dela:

Tekmovanja so sestavni del športa, glavni cilj tekmovanja pa je izbrati najboljšega tekmovalca ali najboljšo ekipo. Poznamo mnogo različnih oblik tekmovanj, ki določajo, kako se tekmovalci pomerijo med seboj: liga, turnir na izpadanje, liga s končnico, ipd... Glavni cilj naloge je implementirati nekaj različnih oblik tekmovanj in s pomočjo simulacij oceniti verjetnost, da najboljša ekipa na koncu tudi dejansko zmaga. Pri tem uporabimo nek smiseln model, ki povezuje izide posameznih dvobojev in moč ekip (npr. model rating ELO).

MENTOR: doc. dr. Erik Štrumbelj

IZJAVA O AVTORSTVU DIPLOMSKEGA DELA

Spodaj podpisani Rok Fujs, z vpisno številko **63070080**, sem avtor diplomskega dela z naslovom:

Ocenjevanje verjetnosti zmage favorita v različnih oblikah tekmovanj

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom doc. dr. Erika Štrumblja,
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".

V Ljubljani, dne 2. oktobra 2014

Podpis avtorja:

Zahvalil bi se mentorju doc. dr. Eriku Štrumblju za usmerjanje in pomoč pri izdelavi diplomske naloge. Zahvalil bi se tudi družini, ki mi je v času celotnega študija stala ob strani in me podpirala.

Kazalo

Povzetek

Abstract

1	Uvod	1
2	Različni tipi tekmovanj	3
2.1	Liga	3
2.2	Turnir na izpadanje	4
2.3	Kombinirano tekmovanje	6
3	Teoretična podlaga	11
3.1	Določanje verjetnosti zmage najboljšega	11
3.2	Velikost napake	12
3.3	Normalna porazdelitev	13
3.4	Rating ELO	14
4	Programska oprema in okolje	17
4.1	Eclipse	17
4.2	Java	18
5	Lastna implementacija tekmovanj	19
5.1	Kreacija ekip	19
5.2	Določanje verjetnosti zmage	21
5.3	Način ocenjevanja tekmovanj	22

KAZALO

5.4	Načini za določanje parov ali skupin	22
5.5	Liga	25
5.6	Turnir na izpadanje	26
5.7	Kombinirano tekmovanje	26
6	Rezultati simulacij	31
6.1	Liga	31
6.2	Turnirji na izpadanje	32
6.3	Kombinirana tekmovanja	33
7	Sklepne ugotovitve in nadaljnje delo	39

Povzetek

Cilj diplomske naloge je bil oceniti verjetnosti favorita v različnih tekmovanjih in podati odgovor na vprašanje, katero tekmovanje je najboljše. Najboljše je tu smatrano kot tisto tekmovanje, ki poda najboljši rezultat z vidika verjetnosti, da bo zmagalo najboljše moštvo. Realizirani so turnirski sistemi liga, turnir na izpadanje in kombinirana tekmovanja. Implementirani so bili tudi različni načini določanja nasprotnikov: žreb, sistem za razvrščanje, sistem za ponovno razvrščanje in kakovostni bobni. Za realizacijo je bil uporabljen programski jezik Java. Na podlagi ustreznega števila zagnanih simulacij vseh turnirskih sistemov z različnimi parametri in ponekod različnimi načini določanja nasprotnikov smo pridobili dovolj dobre ocene, da napovemo verjetnosti posameznih turnirskih sistemov. Izkaže se, da je najboljše tekmovanje liga.

Ključne besede: Favorit, turnir, liga, simulacija.

Abstract

The goal of this thesis was to evaluate the probability of the favourite winning in different competitions and answer the question which competition is the best. The best in this case meaning the competition that gives the best team the highest chance of winning. Tournament systems that were realized are league, elimination tournament, and combined tournaments. Also implemented are different ways of determining opponents. These are draw, seeding, and re-seeding. Java was used as the programming language. On the basis of sufficient number of simulations run, we acquired estimations that are good enough to get a good idea of the probabilities of the favourite winning in various tournament systems. In the end, we conclude that the best competition is the league.

Key words: Favourite, tournament, league, simulation.

Poglavje 1

Uvod

V športu poznamo veliko različnih tekmovalnih sistemov oziroma oblik tekmovalj, ki določajo, kako se bodo tekmovalci in ekipe pomerili med seboj. V vsakem športu vidimo rezultatsko boljše in slabše športnike. Prav nič ni narobe da običajno zmaga boljši nasprotnik in bolj redko slabši. Kaj pa ko zmaga šibkješi nasprotnik? V primeru tekmovalnega sistema imenovanega liga to še ni usodno zaradi narave tega tekmovalja, ki ga bomo opisali v nadaljevanju. Kaj pa ko imamo nek drug turnirski sistem tekmovalja? Vzemimo primer turnirja na izpadanje z eno tekmo. V tem načinu tekmovalja lahko že najmanjša napaka vrže najboljšega tekmovalca iz poti do zmago-slavlja, saj popravnega izpita žal nima. Podrobnejši opis tega tekmovalja bo sledil v naslednjem poglavju. Vseeno pa se lahko čisto iskreno vprašamo, ali je tak način tekmovalja dober za najboljše tekmovalce? Morebiti imajo zato v ligah kot so NBA¹ ali pa NHL² v sklepnih delih tekmovalja turnir na izpadanje, ki določi zmagovaca po štirih zmagah oziroma porazih. Mogoče bi bil ta sistem boljši od prejšnjega in bi dal večjo možnost tekmovalcu ki velja za favorita.

Kateri tip turnirskega tekmovalja je tisti, ki je najbolj ugoden za tekmo-

¹NBA (kratica za National Basketball Association) je zveza moške profesionalne košarke v ZDA[1].

²NHL (kratica za National Hockey League) je profesionalna liga v hokeju na ledu, ki jo sestavlja 30 moštev iz Severne Amerike[2].

valca, ki velja za favorita? Radi bi poiskali tip tekmovanja, ki bo dal največjo možnost zmage favoritu. V nadaljevanju si bomo pogledali oblike tekmovanj kot so liga in turnir na izpadanje ter različne variacije kombinacij teh dveh tipov. Omenili bomo tudi pojme kot so sistem za razvrščanje, naključni žreb in žreb s kakovostnimi razredi. V nadaljevanju bomo namesto pojma sistem za razvrščanje uporabljali termin seeding. Seeding in oba načina žreba bomo poskušali priključiti različnim tipom tekmovanj, bolj natančno turnirjem na izpadanje in kombiniranim tekmovanjem. Predstavili bomo tudi vsa različna tekmovanja in lastne implementacije teh tekmovanj. Podali bomo tudi način ocenjevanja uspešnosti posameznih tekmovanj, kreacijo različnih tekmovalcev oziroma bolje rečeno ekip ter seveda potrebno teorijo za naše delo. Na koncu bomo primerjali implementirane sisteme tekmovanj med seboj in poskušali podati odgovor na zastavljeno vprašanje, kje bo najbolj favorit najbolj uspešen.

Poglavje 2

Različni tipi tekmovanj

2.1 Liga

Liga [3] je tekmovanje, kjer več športnih ekip ali posameznikov tekmuje med seboj v določenem športu. V najpreprostejši obliki je to lahko lokalna skupina amaterjev, ki med seboj oblikujejo ekipe, katere nato tekmujejo med seboj. V najkompleksnejši obliki pa je to lahko mednarodna profesionalna liga, ki služi velike količine denarja in vsebuje na ducate ekip in na tisoče ljudi oziroma tekmovalcev. Sama struktura lige omogoča tekmovanje v določenem razporedu, čemur pogosto pravimo sezona. Rezultat posamezne sezone lige nam poda zmagovalca. Ligaški način tekmovanja zagotavlja da se tekmuje po načinu imenovan krožni sistem. Krožni sistem je sistem, kjer se vsaka ekipa pomeri z vsemi drugimi ekipami. Število tekem med dvema ekipama je ponavadi sestavljeno iz polovice tekem na domačem igrišču in polovice tekem na nasprotnikovem igrišču. Faktor prednosti domačega igrišča je pogost v velikem številu športov. Prednost domačega igrišča je lahko zelo subjektivna. V večini primerov se pokaže v navijačih (večje število na domačih tekmah), v samem domačem igrišču (igranje v bolj poznanem okolju), v potovanju do lokacije (daljša potovanja vplivajo na pripravljenost) ali pa čisto nekaj drugega. V takem tipu tekmovanja je zmagovalec tisti, ki ima najboljši rezultat. Rezultat lahko temelji na:

1. Strogemu sistemu zmaga-poraz-remi.
2. Točkovnemu sistemu, kjer se dodeli določeno število točk za zmago, poraz ali remi. V nekaterih športih se lahko dodelijo dodatne točke, če ekipa doseže določene kriterije.

2.2 Turnir na izpadanje

Turnirju na izpadanje se lahko imenuje tudi pokalni sistem ali izločilni sistem. Uporablja se za določitev zmagovalca na turnirjih v športnih tekmovanjih. Ena ekipa se pomeri z drugo. Zmagovalna ekipa napreduje, medtem ko je poražena ekipa izločena iz tekmovanja. Tak način tekmovanja omogoča hitro določitev zmagovalca, saj se po vsakem krogu turnirja število ekip razpolovi. Slabost pokalnega sistema je v tem, da mora končni zmagovalec premagati vse svoje nasprotnike. Število tekmovalnih ekip mora biti potenca števila dve. Če je preveč ekip, se izvedejo kvalifikacije [4].

Primer: imamo deset ekip. Potenca števila dva, ki je manjša in hkrati najbližja številu deset je osem. Da dobimo samo osem ekip, lahko z žrebom določimo štiri ekipe, ki se bodo med seboj pomerile. Zmagovalni ekipi se pridružita ostalim šestim v turnirju. Drug način kvalifikacij bi lahko bilo tekmovanje ki smo ga malo prej opisali kot liga. Ekipe, ki bi končale ligo na mestih ena do osem, bi se udeležile turnirja.

Kdo se bo pomeril s kom v turnirju lahko določimo na več načinov. Najbolj preprost način je žreb. Če pa so kvalifikacije potekale kot liga ali pa če imamo nek drug način ovrednotejnja ekip, pa lahko izvedemo seeding. Poznamo več vrst seeding-a:

1. začetni seeding,
2. seeding v vsakem krogu,
3. seeding v določenih krogih.

Začetni seeding [5] je prikazan na sliki 2.1, kjer vidimo kako bi potekal seeding šestnajstih ovrednotenih kvalifikantov. Prikazuje tudi nadaljnji potek tekmovanja, kjer opazimo, da za konkretni primer vedno zmaga višje uvrščeni kvalifikant. Ta vrsta seeding-a se izvede samo enkrat in sicer v prvi fazi (na sliki pod oznako first round), ko določamo, kdo se bo pomeril s kom. Najbolje ovrednoteni se pomeri z najšibkejšim izmed kvalifikantov (splošno 1:N, v tem primeru 1:16), drugi po ovrednotenju proti drugemu najslabšemu kvalifikantu (splošno 2:N-1, konkretno v primeru 2:15) itd. Seveda je smiselno dobljene pare tudi pametno razvrstiti, da ne pride že v drugem krogu do kakšne nesmiselnosti, kot bi bil par 1:2 ali pa 1:3. Da se izognemo takim scenarijem, je pametno začeti pri samo dveh tekmovalcih in upoštevati prej omenjeno pravilo 1:N. V naslednjem koraku bi podvojili število tekmovalcev in zopet upoštevati pravilo 1:N pri tem pa bi upoštevali še prejšnji vrstni red tekmovalcev, ki so prej sestavljali pare in vmes vrinili nove tekmovalce ter na ta način tvorili nove pare. Postopek ponavljamo, dokler ne dosežemo števila kvalifikantov. Če se zopet osredotočimo na sliko 2.1 začnemo z oznako finals in vidimo 1:2. Nadaljujemo z oznako semifinals, kjer se nam število tekmovalcev podvoji. Tu iz 1:2 ustvarimo 1:4 in 2:3. Popolnoma vseeno je, ali pišemo 2:3 ali 3:2, kot je na sliki, saj je končni rezultat enak (pomembni so pari v vsakem koraku in ti ostanejo isti). Postopek ponovimo še za oznaki second round in first round da dobimo začetni seeding. Nadaljevanje tekmovanja poteka po principu združevanje sosednjih zmagovalcev v pare, kot je razvidno iz slike 2.1.

Na sliki 2.2 vidimo seeding v vsakem krogu pri zaključnem delu tekmovanja NHL. Tu je razlika med samim začetnim seeding-om in seeding-om v vsakem koraku bolj jasna. Seveda je potrebno sprva izvesti začetni seeding, ki je lahko enak prej opisanemu, vendar pa to ni potrebno. Za seeding v vsakem koraku je pomembno upoštevati samo pravilo 1:N in ne tudi razvrstitve samih parov. Roza obarvani del predstavlja eno divizijo oziroma konferenco in modri del drugo. V modrem delu se vidi posebnost. Če napreduje slabše ovrednotena ekipa, se ponovno izvede seeding. Konkretno na modrem delu

slike vidimo da iz prvega kroga napredujejo moštva 8, 7, 3 in 4. Seeding se ponovno izvede in dobimo para 3:8 in 4:7. Tako ima najboljše ovrednotena ekipa vedno nasproti najslabšega iz svoje divizije in s tem tudi večjo možnost za zmago. Če bi imeli samo začetni seeding, bi v tem primeru dobili para 8:4 in 3:7 (za lažje razumevanje si pogledjmo sliko 2.1 pod oznako second round in si zamislimo, da v naslednji krog označen z semifinals napredujejo 8, 7, 3 in 4). Seeding v določenih krogih pa bi se izvajal v poljubnem oziroma poljubnih krogih.

Načini napredovanja v turnirju na izpadanje so:

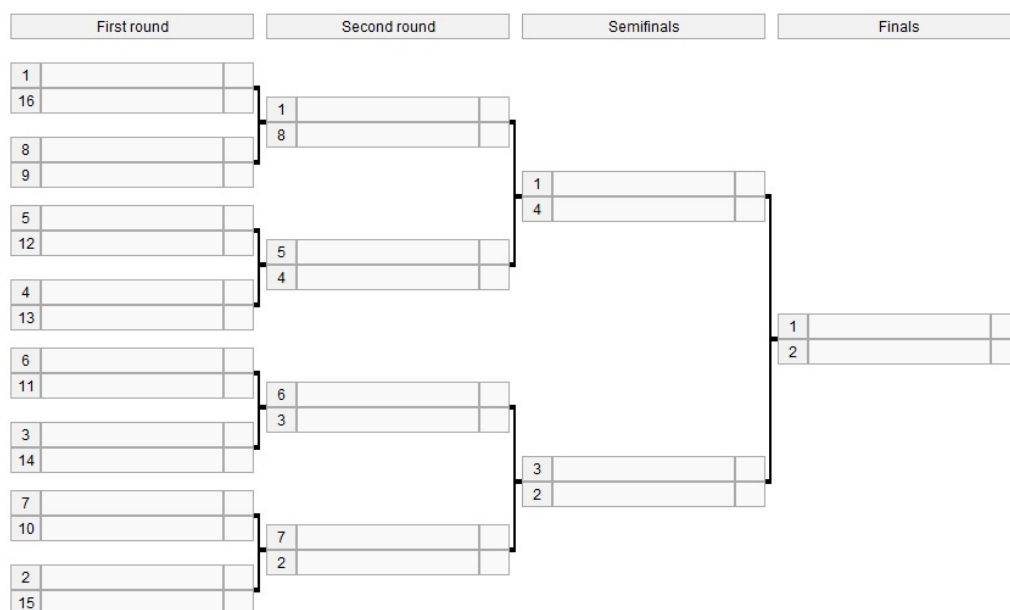
1. Zmaga na en dvoboj,
2. zmaga na dva ali več dvobojev in
3. v nekaterih športih tudi zmaga na seštevek rezultata dveh ali več dvobojev.

Tekmovanje se konča, ko se izvede še finale (imamo samo še dva tekmovalca), katerega zmagovalec je tudi zmagovalec turnirja na izpadanje.

2.3 Kombinirano tekmovanje

Kot predlaga že samo ime, je kombinirano tekmovanje sestavljeno iz več različnih načinov tekmovanj.

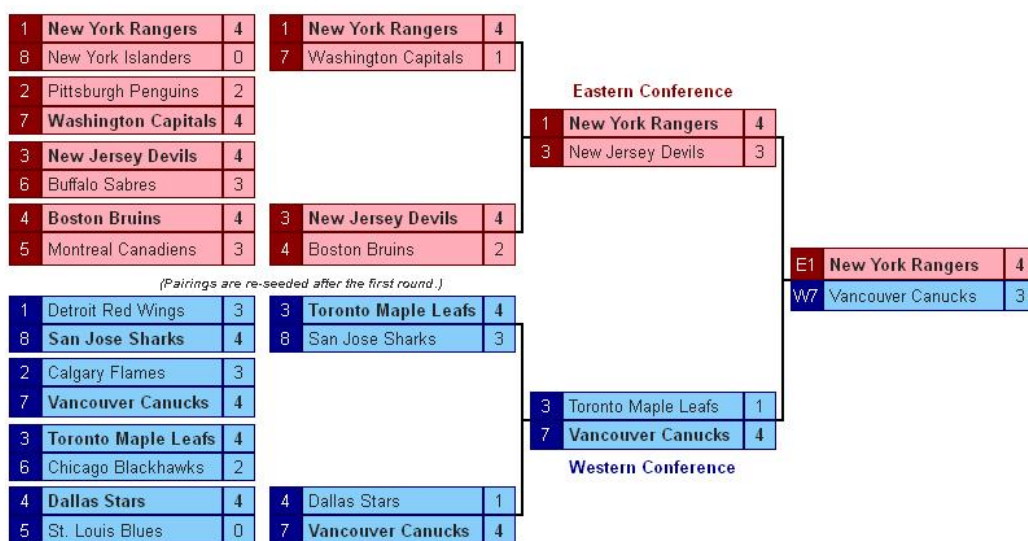
Za lažjo predstavitev vzemimo primer sistema tekmovanja svetovnega prvenstva v nogometu[6]. Na to tekmovanje se uvrsti 32 ekip. Razdeljene so v skupine velikosti štiri, torej osem skupin. V katero skupino spada katera ekipa, določi žreb iz kakovostnih bobnov (angleško pots)[7]. Nato se v skupinah vsaka ekipa enkrat pomeri z vsemi v skupini (krožni sistem). Ta del tekmovanja bi sedaj lahko označili z liga, kjer ima samo gostitelj ali gostitelji prednost domačega igrišča. Za ostale pare udeležencev, ki se pomerijo med seboj in v paru ni gostitelja, se kraj tekmovanja označi kot nevtralen. Ko so se pomerile vse ekipe znotraj skupine (vsaka ekipa po tri tekme) in



Slika 2.1: Začetni seeding kvalifikantov in nadaljnji potek tekmovanja.

so se za rezultat dodelile točke, dobimo vrstni red. Prvi dve ekipi skupine napredujeta v zaključni del tekmovanja, preostali dve sta zaključili s tekmovanjem. Izvede se križanje med skupinami. Križata se sosednji skupini. Če vzamemo za primer skupino A in skupino B, lahko označimo zmagovalca skupine A kot A1 in A2, ter konsistentno tudi zmagovalca skupine B kot B1 in B2. Zelo intuitivno je da se bo A1 pomeril z B2 in tudi da se bo B1 pomeril z A2. Torej v splošnem močnejši iz ene skupine z šibkejšim iz druge skupine. Tudi to lahko označimo z pojmom seeding, le da imamo sedaj predpogoj A se križa B, nato pa se zopet uporabi pravilo 1:N iz seeding. Za ilustracijo vzemimo sliko 2.3, ki prikazuje primer s letošnjega svetovnega prvenstva. Po končanem skupinskem delu tekmovanja so se tako od tekmovanja poslovile reprezentance uverščene na 3. in 4. mestu obeh skupin. Pot v osmino finala pa so nadljevale reprezentanca Brazilije proti Čilu (A1:B2) in reprezentanca Mehike proti Nizozemski (A2:B1).

Ta del tekmovanja je sedaj turnir na izpadanje, potek katerega smo že



Slika 2.2: Seeding v vsakem koraku.

opisali. Kvalifikantje tega dela tekmovanja, turnirja na izpadanje, pa so A1, A2, B1, B2, C1, C2, D1, D2, E1, E2, F1, F2, G1, G2, H1 in H2. Potek tekmovanja od tu naprej je enako kot turnir na izpadanje, kjer se igra na eno zmago za napredovanje. Omeniti velja, da se na takih velikih tekmovanjih pogosto igra tudi za tretje mesto, torej poraženca polfinal se pomerita za tretje mesto. Temu se pogosto pravi mali finale.

Standings		Over/Under	Top Scorers							
#	▲ Group A	MP	W	D	L	G	Pts	Form		
1.	Brazil	3	2	1	0	7:2	7	W	D	W
2.	Mexico	3	2	1	0	4:1	7	W	D	W
3.	Croatia	3	1	0	2	6:6	3	L	W	L
4.	Cameroon	3	0	0	3	1:9	0	L	L	L
#	▲ Group B	MP	W	D	L	G	Pts	Form		
1.	Netherlands	3	3	0	0	10:3	9	W	W	W
2.	Chile	3	2	0	1	5:3	6	L	W	W
3.	Spain	3	1	0	2	4:7	3	W	L	L
4.	Australia	3	0	0	3	3:9	0	L	L	L

Slika 2.3: Primer s svetovnega prvenstva v nogometu 2014.

Poglavje 3

Teoretična podlaga

Da podamo odgovore na vprašanja diplomske naloge, bo potrebno znati izračunati verjetnost, da zmaga favorit na določenem tekmovanju, z določenim številom nasprotnikov, z različnimi kvalitetskimi razlikami med tekmovalci, ipd...

3.1 Določanje verjetnosti zmage najboljšega

Sistem tekmovanja, različno število nasprotnikov ter njihova kvaliteta bodo za nas parametri, ki jih lahko označimo z nekim skupnim simbolom θ .

Vsako tekmovanje je slučajen proces. Zaradi tega imamo pri določenih parametrih θ več različnih scenarijev oziroma potekov, kako se posamezno tekmovanje lahko odvije (kdo je zmagovalec neke tekme, kakšen je seeding, ali je naključen, ipd....).

Vsak scenarij $x \in \Omega$ se bo zgodil z neko določeno verjetnostjo $p(x)$. Z Ω sedaj označimo množico vseh možnih scenarijev.

Ponovimo do sedaj napisano s simboli:

$$P(\text{Zmaga_najboljši}|\theta) = \int I(\theta, x)p(x)dx, \quad (3.1)$$

kjer je $I(\theta, x)$ preprosta funkcija, ki zavzame vrednost 1, če je favorit zmagal v scenariju x , sicer pa zavzame vrednost 0. Ta enačba nam na formalen

način pove, da verjetnost zmage favorita dobimo tako, da se sprehodimo skozi vse možne scenarije in pogledamo, ali tam zmaga ali ne, pri tem pa vsak scenarij utežimo z njegovo verjetnostjo.

Integral 3.1 pa ni tako enostavno izračunati, kajti scenarijev je veliko (ob upoštevanju različnih moči ekip, jih je celo neskončno). Prav tako ni enostaven izračun verjetnosti posameznega scenarija $p(x)$. Edina stvar, ki jo zagotovo znamo izračunati, je $I(\theta, x)$. Če imamo podan določen scenarij in vse parametre tekmovanja, potem lahko samo pogledamo, ali je favorit zmagal.

Zaradi teh razlogov se zatečemo k simulaciji, natančneje, metodi Monte Carlo. Namesto, da bi verjetnost izračunali, jo bomo zgolj ocenili na podlagi slučajnega vzorca scenarijev x_1, x_2, \dots, x_n . Po metodi Monte Carlo je

$$P(\text{Zmaga najboljši}|\theta) \approx \frac{1}{n} \sum_{i=1}^n I(\theta, x_i) \quad (3.2)$$

To je formalni zapis poglobitnega pri našem delu. Generirali smo nekaj naključnih scenarijev, prešteli kolikokrat je zmagal favorit in na koncu vzeli delež scenarijev, v katerih je favorit zmagal, kot oceno verjetnosti njegove zmage.

Postopek prehoda iz enačbe 3.2 v enačbo 3.1 se imenuje tudi integriranje z metodo Monte Carlo (angl. Monte Carlo Integration). Osnova je zakon velikih števil [8]. Ta zakon pa v osnovi pravi, da se povprečje slučajnega vzorca bliža upanju porazdelitve, ko povečujemo velikost vzorca, ter da bo skoraj gotovo slej ali prej poljubno natančen.

3.2 Velikost napake

Več vzorcev kot bomo vzeli, manjša bo naša napaka. Natančneje, standardni odklon ocene je $\frac{\sigma_I}{\sqrt{n}}$, kjer je σ_I standardni odklon zmage favorita. Ta standardni odklon je v našem primeru $\sqrt{p(1-p)}$, kjer je p verjetnost zmage favorita. Do teh $\sqrt{p(1-p)}$ smo prišli, ker je to standardni odklon binarne/binomske slučajne spremenljivke. Vemo, da je I taka spremenljivka, saj

zavzame le vrednosti 1 ali 0 z določeno verjetnostjo. Osnova za izpeljavo napake, ki jo naredimo, če ocenjujemo s povprečjem vzorca je centralni limitni izrek (CLI). CLI pravi, da ko je vzorec dovolj velik, bo povprečje vzorca porazdeljeno približno normalno s standardnim odklonom, ki bo za koren od velikosti vzorca manjši od standardnega odklona originalne porazdelitve, iz katere vzorčimo[8]. Torej, napaka pada sorazmerno s korenom velikosti vzorca. V našem konkretnem primeru je standardni odklon originalne porazdelitve $\sqrt{p(1-p)}$, napako pa dobimo, če še delimo s korenom velikosti vzorca $\frac{\sigma_L}{\sqrt{n}}$.

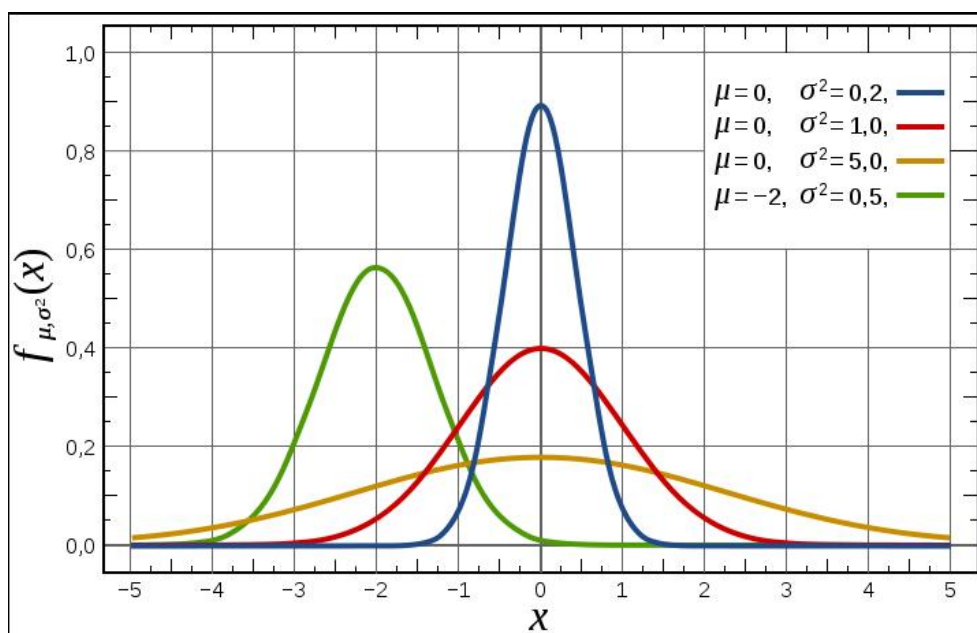
3.3 Normalna porazdelitev

Normalna porazdelitev, poznana tudi kot Gaussova porazdelitev, je verjetnostna porazdelitev vrednosti statističnih enot v statistični populaciji, ki je v grafični predstavitvi oblikovana v obliki zvona oziroma normalne krivulje. Gre za družino porazdelitev z dvema parametroma (aritmetična sredina in standardni odklon). Standardizirana normalna porazdelitev je porazdelitev, ki zavzema vrednosti s standardnim odklonom 1 in aritmetično sredino (povprečjem) 0.

Normalna porazdelitev je izrednega pomena za kvantitativne metode različnih znanosti. Po normalni krivulji se porazdeljuje človekova višina in masa, stopnja IQ, itd... Mi pa bomo predpostavili, da se tudi moč ekip v nekem tekmovanju porazdeljuje približno normalno, a z neznanim standardnim odklonom (z drugimi besedami, dopuščamo možnost, da so ekipe bolj ali pa manj izenačene po moči).

Predpostavljanje normalne porazdelitve je bistveno za množico statističnih izračunov, saj po centralnem limitnem izreku velja, da se vzorec, ki je izvzet iz celotne populacije, porazdeljuje približno po normalni krivulji tudi, če vrednosti vseh enot matične populacije niso porazdeljene normalno[9].

Nekaj primerov normalne porazdelitve lahko vidimo na sliki 3.1



Slika 3.1: Funkcije gostote verjetnosti za normalno porazdelitev.

3.4 Rating ELO

Eksperimenti, ki smo si jih zastavili, od nas zahtevajo tudi simulacije tekmovalj z različnimi ekipami glede na to, kako izenačene so med seboj po moči. V ta namen moramo definirati moč ekipe in model, ki povezuje moč dveh ekip z verjetnostjo izida dvoboja med njima. Uporabili bomo model, ki se skriva za ratingom ELO.

Rating ELO [10] je bil sprva način ocenjevanja sposobnosti igralcev šaha, sedaj pa se uporablja tudi kot način ocenjevanja za razne video igre in ekipne športe kot so nogomet in košarka. Rating ELO je dobil ime po svojem kreatorju Arpadu El [11].

Razlika med ratingoma ELO služi kot napoved dvoboja med dvema ekipama. Dve ekipi z približno enako močjo bosta imeli približno enako število zmag (50%-50%). Če ima na primer neka ekipa oceno moči večjo za 100 od druge ekipe, je pričakovan rezultat prve ekipe nekje 64%, pri razliki 200 pa okoli 76%. Moč ekipe je predstavljena z številom, ki se povečuje oziroma

zmanjšuje gleda na izide posameznih odigranih dvobojev. Po vsakem dvo-
boju zmagovalna ekipa odvzame točke poraženki. Sama razlika v moči po
ELO pa določa število točk dodeljenih ozrioma odvzetih za po končani tekmi.
To pa pomeni da se ELO sam popravlja. Če ima neka ekipa A oceno moči
 R_A in ekipa B oceno moči R_B , potem po ELO zapišemo verjetnost zmage
ekipe A proti ekipi B kot

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}}, \quad (3.3)$$

in verjetnost zmage ekipe B proti A kot

$$E_B = \frac{1}{1 + 10^{(R_A - R_B)/400}}. \quad (3.4)$$

Če vzamemo malce drugačen zapis pa lahko zgornji enačbi izrazimo tudi kot

$$E_A = \frac{Q_A}{Q_A + Q_B}, E_B = \frac{Q_B}{Q_A + Q_B}, \quad (3.5)$$

kjer je $Q_A = 10^{R_A/400}$ in $Q_B = 10^{R_B/400}$. V drugi notaciji lahko opazimo,
da imata Q_A in Q_B enak imenovalec. To pa pomeni da za vsakih 400 razlike v
moči med ekipama se verjetnost zmage ene ekipe poveča za faktor 10. Dokaz
za slednjo trditev:

$$\begin{aligned} E_A &= X E_B \\ \frac{Q_A}{Q_A + Q_B} &= X \frac{Q_B}{Q_A + Q_B} \\ Q_A &= X Q_B \\ 10^{R_A/400} &= X 10^{R_B/400} \\ \frac{10^{R_A/400}}{10^{R_B/400}} &= X \\ 10^{R_A/400 - R_B/400} &= X \\ 10^{\frac{1}{400}(R_A - R_B)} &= X, \end{aligned}$$

kjer smo rekli, da je absolutna razlika med R_A in R_B enaka 400, zato sledi

$$10 = X$$

Zabeležimo še, da je $E_A + E_B = 1$. Ker pa v praksi prava moč ekipe ni znana, se pričakovan izid dvoboja izračuna preko trenutnih moči ekip. Če ekipa presega pričakovane rezultate, ELO to vzame kot dokaz, da je trenutna moč ekipe prenizka in se mora povečati. Podobno velja za obratni primer, ko ekipa ne dosega pričakovanih rezultatov je potrebno njeno moč zmanjšati. Arpad Elo je tu predlagal način, ki se še vedno močno uporablja. Njegov način je majhna linearna sprememba, ki je proporcionalna številu tekem, ki jih je ekipa odigrala nad ali pa pod pričakovanji ELO. Največja možna sprememba, imenovana K-faktor, je bila postavljena na $K=16$ za boljše in $K=32$ za slabše ekipe. Predpostavljajmo da je bilo pričakovano od ekipe A da doseže H_A točk. Dejansko pa je dosegla S_A točk. Formula za posodobitev ekipe A bi bila

$$R'_A = R_A + K(S_A - H_A). \quad (3.6)$$

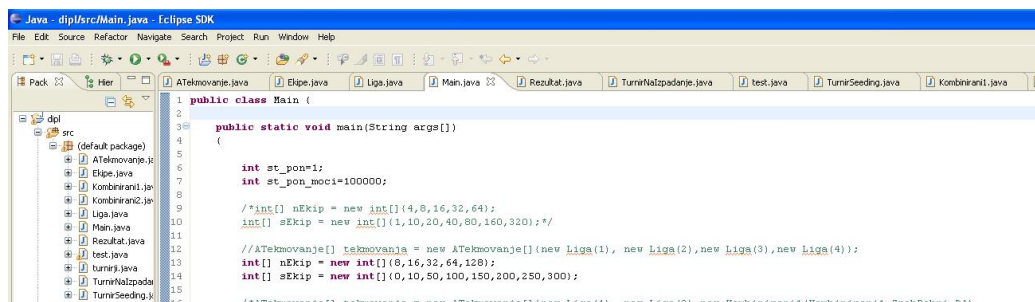
Več o ELO si lahko pogledamo na [10]

Poglavje 4

Programska oprema in okolje

4.1 Eclipse

Za programiranje kode smo uporabili programsko orodje Eclipse[12]. Eclipse je integrirano razvijalno okolje, ki primarno uporablja jezik programski Java za razvoj aplikacij. Izgled Eclipse je se vidi na sliki 4.1. Preko raznih dodatkov pa se lahko razvija aplikacije v drugih programskih jezikih. Ti jeziki so Ada, ABAP, C, C++, COBOL, Fortran, Haskell, JavaScript, Lasso, Natural, Perl, PHP, Prolog, Python, R, Ruby, itd...



Slika 4.1: Orodje Eclipse.

4.2 Java

Za programski jezik smo si izbrali programski jezik Java[13]. Java je objektno orientiran jezik. Glavna prednost jezika je omogočanje kodi, ki jo napiše razvijalec, da se prevedena lahko izvaja na več operacijskih sistemih (OS), ne da bi jo bilo potrebno ponovno prevajati za drug OS.

Poglavje 5

Lastna implementacija tekmovanj

Za kreiranje moči oziroma ratingov ekip smo uporabili normalno porazdelitev (poznana tudi kot Gaussova porazdelitev) s parametri aritmetična sredina (M) in standardni odklon (S) kot je prikazano na tabeli 5.1. Seveda pa smo to počeli za različno število ekip (N).

5.1 Kreacija ekip

parametri	sopomenka v kodi	vrednost
M	popvprecnaMoc	1500
S	standardniOdklon	0 - 300
N	n	8 - 128

Tabela 5.1: Tabela parametrov in njihove vrednosti

Parameter M ne igra bistvene vloge, vrednost 1500 pa smo izbrali zato, da bodo moči ekip podobne ratingom pri šahu. Parameter S (standardni odklon) lahko interpretiramo kot mero izenačenosti med ekipami. Torej če je $S=0$, potem vemo, da so vse ekipe enako močne, če je S majhen so si ekipe

zelo podobne v moči in če je S velik, potem so razlike v moči velike. Priložili smo tudi realizacijo kreacij ekip v obliki kode 5.1.

Podajmo še primer klica kode za kreacijo osmih ekip s parametri $M=1500$ in $S=200$. Tu imamo še parameter `randomSeed`, ki spreminja začetni seed naključnega generatorja. Tu smo se odločili, da uporabimo nekaj preprostega, kot je čas. Torej klic kode 5.1 bi izgledal kot prikazuje 5.3.

Listing 5.1: Java koda za kreiranje ekip

```
public static double[] kreiraj_ekipe(double povprecnaMoc, double
    standardniOdklon, int n, long randomSeed)
    // dodamo random seed, da lahko fiksiramo eksperiment, ce
    zelimo
    {
        //kreiraj ekipe
        double[] ekipe=new double[n];
        java.util.Random random1 = new
            java.util.Random(randomSeed);
        for(int i=0;i<n;i++){
            ekipe[i]=povprecnaMoc+
                standardniOdklon*random1.nextGaussian();
        }
        return ekipe;
    }
```

Listing 5.2: Primer klica kode za kreiranje ekip

```
double[] mocEkip =
    Ekipe.kreiraj_ekipe(1500,s,n,System.currentTimeMillis());
```

5.2 Določanje verjetnosti zmage

Tu smo se odločili za pristop po ELO, ki je bil opisan že v poglavju 3.5. Razlog za to leži v tem, da je ELO relativno enostaven model, ki pa obenem precej dobro opisuje dejansko dogajanje v športu. Predpostavili smo, da so generirane moči ekip tudi prave moči ekip, ki se med samim tekmovanjem ne spreminjajo.

Torej v izseku kode 5.3 je prikazano kako smo implementirali enačbi 3.5, ki se nahajata v poglavju 3.5. Edina sprememba je, da smo dodali še spremenljivko `prednostDomacina`, s katero lahko vplivamo na faktor prednosti domačega igrišča. V primeru, ko je ta spremenljivka postavljena na 0, je koda, ki se izvede, enakovredna prej omenjenima enačbama.

Listing 5.3: Izračun verjetnosti zmage ekip

```
public static double[][] verjetnosti_zmage(double[] mociEkip,
    double pD){
    double[][] verjetnosti=new
        double[mociEkip.length][mociEkip.length];
    for (int i=0;i<verjetnosti.length;i++){
        for(int j=0;j<verjetnosti[0].length;j++){
            double u_Domacin =
                Math.pow(10,((mociEkip[i]+pD)/400));
            double u_Gost =
                Math.pow(10,((mociEkip[j]-pD)/400));
            if(i!=j) verjetnosti[i][j]= u_Domacin
                / (u_Gost + u_Domacin);
        }
    }
    return verjetnosti;
}
```

Funkcija v kodi 5.3 nam torej vrne tabelo verjetnosti po tem, ko smo ji podali parametre moči ekip in prednost domačina. Omeniti velja, da pred-

nosti domačina p_D v naših poskusih nismo uporabljali, ampak nastavili na 0.

5.3 Način ocenjevanja tekmovanj

Za določanje verjetnosti zmage favorita posameznega tekmovanja pri podanih parametrih N in S je bilo potrebno izvesti veliko število simulacij. Ustavili smo se pri številu 100.000 simulacij. To številko je dovolj za naše verjetnosti, saj je velikost naše napake enaka $\sqrt{n^{-1}}\sqrt{p(1-p)}$ (glej poglavje 3.3). Ta bo dosegla največjo vrednost, če bo $p = 0.5$. Koren iz sto tisoč pa znaša okoli 316. Torej če vstavimo vrednosti, dobimo $\frac{\sqrt{0.25}}{\sqrt{100000}} \approx 0.0016$. To pomeni da je naša napaka manjša od 0.2%. S tem so naše ocene že dovolj natančne, da jih lahko uporabimo za rezultat.

Simulacija je ovrednotena z 1, ko favorit zmaga, in 0, ko ne zmaga. Iz generiranih moči vedno vemo, kdo je favorit.

5.4 Načini za določanje parov ali skupin

5.4.1 Navaden žreb

Žreb lahko uporabimo tako pri določanju parov v turnirjih na izločanje kot pri sestavljanju skupin v skupinskih delih kombiniranih tekmovanj. Zamislili smo si način, kjer funkcija prikazana na 5.4 sprejme indekse udeležencev nekega tekmovanja, jih naključno premeša ter vrne pomešani vrstni red.

Listing 5.4: Navaden žreb

```
public static int[] zreb(int[] array){
    int index;
    int temp;

    java.util.Random random = new java.util.Random();
    for (int i = array.length-1;i>0;i--){
```

```
        index = random.nextInt(i+1);

        if (index != i){
            temp=array[index];
            array[index]=array[i];
            array[i]=temp;
        }
    }
    return array;
}
```

5.4.2 Začetni seeding

Začetni seeding se drži pravila najmočnejši proti najšibkejšemu (1:N, 2:N-1,..., M:(N+1-M)). Upoštevati mora pa tudi razvrstitev dobljenih parov. Ker je celotna koda za začetni seeding zapletena in dolga, bomo tukaj opisali samo ključni del za začetni seeding. To je pridobitev samih indeksov parov v pravilnem vrstnem redu.

Torej, v kodi 5.5 je prikazan začetni seeding. Funkcija `seeding2` kot parametere prejme dve tabeli integerjev. Kot vidimo je funkcija rekurzivna in se ustavi, ko sta prejeta parametra enaka v velikosti. Omenimo še, da to ni celotna funkcija, ki se uporablja v diplomski nalogi, ampak služi kot v oporo drugi malce večji funkciji. Vendar pa vseeno lahko že iz `seeding2` dobimo željene pravilne pare. Ključ do tega je v pravilnem klicu, ki je prikazan s kodo 5.6.

Listing 5.5: Začetni seeding

```
public static int[] seeding2(int[] ekipe,int[] start){
    int[] seating = new int[start.length*2];
    for(int i = 0; i<seating.length; i += 2){
        seating[i] = start[i/2];
        seating[i+1] = seating.length-1-seating[i];
    }
}
```

```
    }
    if(ekipe.length==seating.length){
        return seating;
    }
    else{
        return seeding2(ekipe,seating);
    }
}
```

Listing 5.6: klic funkcije5.5 in njen izpis

```
int[] ekipe={1,2,3,4,7,10,13,20};
int[] start={0};
    for (int i:TurnirSeeding.seeding2(ekipe, start)){
        System.out.println(i);
    }
*****izpis*****
0
7
3
4
1
6
2
5
```

V javi potekajo indeksi od 0 dalje, zato smo tudi tu prevzeli, da je 0 indeks najmočnejšega tekmovalca. Če kličemo `seeding2` z prvim parametrom, željenim številom udeležencev (elementi niso pomembni, le dolžina seznama), in drugim parameterom, tabelo integerjev z elementom v vrednosti 0, potem vedno dobimo pravilen vrstni red in pravilne pare. Nekoliko nerodno implementirano, vendar omenimo, da ta klic služi zgolj kot iniciacija indeksov neki drugi funkciji in ga uporabniku ni potrebno uporabljati. In, kot vidimo na

izpisu v 5.6, so pravilni tako pari kot njihov vrstni red (le da je zamaknjeno za -1). To pa bo zato ker začnemo pri samem vrhu. 0 je naš najmočnejši tekmovalc in njemu dodamo nasprotnika 1. Nato ponovno kličemo seeding2 in iz para 0:1 dobimo dva nova para 0:3 in 1:2, nato ponovno kličemo seeding2 in iz zadnjih dveh parov dobimo 0:7, 3:4, 1:6 in 2:5.

5.4.3 Seeding v vsakem koraku

Začnemo kar z začetnim seedingom. V vseh nadaljnjih korakih pa vedno upoštevamo pravilo najmočnejši, ki je še v tekmovanju, se bo pomeril proti najšibkejšem, ki je še v tekmovanju.

5.4.4 Kakovostni bobni

Se uporabljajo kot alternativa navadnemu žrebu. Ekipe so najprej ovrednotene po moči. Pogledamo število skupin, ki so v tekmovanju. To število nam pove koliko ekip bo pripadalo nekemu bobnu. Število bobnov je določeno z $B = \frac{N}{\check{S}st}$, kjer je N predstavlja število ekip in $\check{S}st$ pomeni število skupin v tem delu tekmovanja. Tako lahko prvih B ekip uvrstimo v prvi kakovostni bobnen, naslednjih B ekip v drug kakovostni bobnen, itd.

5.5 Liga

Ligo smo implementirali kot beleženje zmag oziroma porazov skozi določeno število dvokol¹. Na koncu lige je zmagovalec tisti, ki zbere največ točk. Implementirali smo strogi način poraz-zmaga. V tem primeru zmaga prinese eno točko, medtem ko poraz prinese nič točk. Pri štetju uspešnih simulacij je bilo potrebno upoštevati primer izenačenj na prvem mestu. Če je bilo pet

¹dvokolo: Dva združena dvoboja, ko se pomerita dve ekipi. En dvoboj predstavlja prvo ekipo kot gostiteljico in drugo kot gostjo. Drugi dvoboj pa pa ekipi zamenjata vlogi gosta in gostitelja.

zmagovalcev z istim številom točk, se šteje, kot da je favorit zmagal v eni petini primerov.

5.6 Turnir na izpadanje

Ta tip tekmovanja smo sprogramirali kot rekurzivno funkcijo, ki sprejeme vse kvalifikante turnirja na izpadanje. Funkcija izvede turnir na izpadanje za podane kvalifikante in nato kliče sama sebe. Sedaj so kvalifikantje v primeru začetnega seeding kar zmagovalci posameznih dvobojev. V primeru seeding v vsakem koraku, pa moramo ponovno klicati funkcijo seeding, ki nam potem tudi vrne pravilno razporeditev preostalih v tekmovanju. To počnemo dokler nam funkcija ne vrne samo še zadnjega zmagovalca, ki je posledično tudi naš zmagovalec turnirja. Implementirali smo tako način žreba kot tudi začetni seeding in seeding v vsakem krogu tekmovanja. Dodali smo še možnost izbire koliko posamičnih zmag je potrebnih, da je nekdo v dvoboju dokončno zmagal. Pri štetju uspešnih simulacij ni bilo posebnosti, saj je lahko favorit le zmagal ali pa izgubil.

5.7 Kombinirano tekmovanje

Implementirali smo dve različni tekmovanji, ki jih bomo imenovali kar K1 in K2. Obe sta sestavljeni tako iz skupinskega dela (podobno tekmovanju liga) kot tudi iz turnirja na izpadanje.

5.7.1 K1

Za kombinirano tekmovanje K1 smo si izbrali način tekmovanja kot ga uporabljajo na svetovnem prvenstvu v nogometu. Torej, kvalifikante, ki morajo biti vsaj štirje in hkrati potenca števila dve, razdelimo v skupine po štiri. Realizirali smo dva načina delitve v skupine. Nobeden od njiju žal ni točno tak kot na svetovnem prvenstvu, saj je potrebno upoštevati kvoto kontinentov, mi pa smo želeli posplošiti tekmovanje na več in manj udeležencev in ne

zgolj točno na 32. Prvi je običajen žreb, drugi pa je žreb s kakovstnimi bobni. Slednji je tisti, ki daje res enakovredne skupine. Realiziran je na način, da dobljene kvalifikante razvrstimo po vrsti kot so ovrednoteni po njihovi moči, nato pa jih razdelimo v kakovostne bobne. V prvi boben spadajo ekipe od števila ena do števila različnih skupin. Torej če imamo šestnajst udeležencev, imamo štiri skupine. To bi pomenilo, da gredo v prvi kakovostni boben prve štiri ekipe po razvrstitvi(1-4). Naslednje štiri(5-8) v drugi kakovostni boben, naslednje štiri(9-12) v tretji boben in še zadnje štiri(13-16) v zadnji četrti boben. Nato pa izvedemo žreb. Najprej dodelimo vse ekipe iz prvega bobna vsako v svojo skupino. Nato ponovimo postopek za drugi, tretji in na koncu še četrti boben. Vsaka izmed skupin izvede skupinski del, kjer se vsaka ekipa v skupini po enkrat pomeri z ostalimi znotraj iste skupine. Prvo in drugo mesto v skupini prinašata napredovanje v drugi del tekmovanja. Drugi del tekmovanja je turnir na izpadanje. Uporabimo začetni seeding. Prvi v skupini igra z drugim v sosednji skupini in drugi v skupini odigra s prvim sosednje skupine. Odigra se po eno tekmo da se določi zmagovalca.

5.7.2 K2

Za kombinirano tekmovanje K2 smo se odločili poskusiti nekoliko prirejeno tekmovanje K1.





Ideja izhaja iz evropskega prvenstva v košarki [14]. Tam se kvalificira 24 ekip, ki so nato razdeljene v štiri skupine po šest ekip. Tu se izvede prvi skupinski del. Med sabo odigrajo po krožnem sistemu, torej po enkrat vsaka z vsako v svoji skupini. Zadnje tri ekipe tu končajo tekmovanje. Napredujejo prve tri ekipe po razvrstitvi iz vsake skupine. Po dve sosednji ekipi se združita in sedaj imamo samo še 2 skupini, ki vsebujeta obe po šest ekip. Točke dosežene v prvem skupinskem delu se prenesejo naprej, vendar samo tiste točke, ki so bile dosežene proti ekipam, ki so tudi napredovale. Sedaj se izvede drugi skupinski del. Tokrat se ekipe znova pomerijo med seboj, vendar samo tiste ki se prej še niso. Torej če so napredovale A1, A2 ter A3 iz skupine A in B1, B2 in B3 iz skupine B, potem to pomeni da se vse ekipe z oznako A(1-

3) pomerijo z ekipami, ki nosijo oznako B(1-3). Na koncu napredujejo prve štiri ekipe iz obeh združenih skupin. Mesto v skupini je seveda pomembno, saj določa nasprotnika v nadaljevanju tekmovanja. Tu se zopet uporablja začetni seeding in izvede se turnir na izpadanje. Za lažje razumevanje prvega in drugega skupinskega dela bomo razložili potek teh dveh delov tekmovanja na primeru. V sliki 5.1 vidimo stanje po končanem prvem skupinskem delu. Če se omejimo na skupino D, vidimo da je Italija premagala vse v svoji skupini, kar pomeni da bo naprej nesla maksimalno število zmag (2). Finska je bila premagana zgolj s strani Italije in bo naprej nesla eno zmago, ki jo je dobila proti Grčiji. Grčija pa nadaljuje tekmovanje brez zmag. Zmage se nahajajo pod oznakami W in WO.

#	Group C	MP	W	WO	LO	L	TP	Pts	Form
1.	Spain	5	4	0	0	1	369:269	9	W W W L W
2.	Croatia	5	3	1	0	1	337:341	9	W W W W L
3.	Slovenia	5	3	0	1	1	347:344	8	L W W W W
4.	Czech Republic	5	2	0	0	3	316:339	7	L W L W L
5.	Georgia	5	1	0	0	4	366:394	6	L L L L W
6.	Poland	5	1	0	0	4	329:377	6	W L L L L
#	Group D	MP	W	WO	LO	L	TP	Pts	Form
1.	Italy	5	5	0	0	0	391:339	10	W W W W W
2.	Finland	5	3	1	0	1	358:337	9	W W L W W
3.	Greece	5	3	0	0	2	392:350	8	L L W W W
4.	Sweden	5	1	0	0	4	345:391	6	L L W L L
5.	Turkey	5	1	0	0	4	355:398	6	L W L L L
6.	Russia	5	1	0	1	3	374:400	6	W L L L L

Slika 5.1: Primer prvega skupinskega dela tekmovanja v evropskem prvenstvu košarke 2013.

V drugem skupinskem delu pa se te tri reprezentance pomerijo z reprezentacami Španije, Hrvaške in Slovenije. Skupini C in D se torej združita in tvorita skupino F. Kaj so katere ekipe odnesle s seboj v ta del tekmovanja se lahko vidi v sliki 5.2 pod oznako Form in sicer najbolj desna dva okence. Zeleno okence z črko W pomeni zmaga in rdeče okence z črko L pomeni poraz. Za hitro preverjanje lahko hitro opazimo, da je Italija res prinesla naprej dve zmagi. V 5.1 je že podan vrstni red in točkovni razplet po odigranih preostalih treh tekem vseh reprezentanc. Napredovale so tako reprezentance na mestih 1-4. Nato pa je sledil še turnir na izpadanje na eno dobljeno tekmo.

#	Group F	MP	W	WO	LO	L	TP	Pts	Form
1.	 Croatia	5	2	2	0	1	372:361	9	W W W W L
2.	 Slovenia	5	3	0	1	1	385:379	8	L W W L W
3.	 Italy	5	2	1	0	2	374:357	8	W L L W W
4.	 Spain	5	2	0	1	2	375:339	7	L W L L W
5.	 Finland	5	2	0	0	3	341:385	7	W L L W L
6.	 Greece	5	1	0	1	3	381:407	6	L L W L L

Slika 5.2: Primer drugega skupinskega dela tekmovanja v evropskem prvenstvu košarke 2013

K2 smo realizirali na zelo podoben način. V realizaciji upoštevamo potenco števila dve, zato naše začetne skupine sestavljajo po štiri ekipe. Izvedemo prvi skupinski del, ki je enak prej opisanemu. Tudi napredujejo po tri ekipe. Edina sprememba je, da sedaj tekmovanje konča le najslabše uvrščena ekipa v skupini. Točke se prenesejo, razen tistih doseženih proti zadnje uvrščeni ekipi, ki je končala s tekmovanjem. Sedaj združimo po dve sosednji skupini in tvorimo novo, ki vsebuje po 6 ekip. Od tu naprej je tekmovanje enako. Torej še drugi skupinski del in nato turnir na izpadanje.

Poglavje 6

Rezultati simulacij

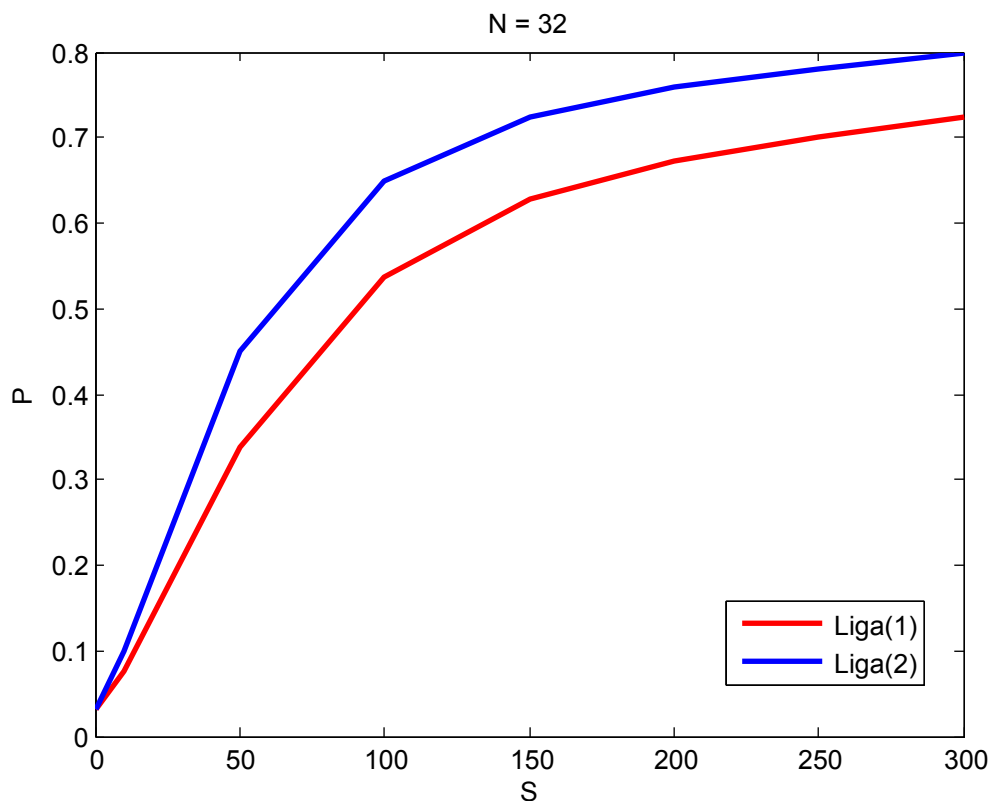
Grafi na slikah v tem poglavju vsebujejo spremenljivke. Pomen spremenljivk je predstavljen v tabeli 6.1.

spremenljivka	pomen
S	standardni odklon
P	verjetnost
N	število ekip

Tabela 6.1: Tabela spremenljivk in njihov pomen

6.1 Liga

Iz grafa na sliki 6.1 vidimo, kako se verjetnost zmage favorizirane ekipe spreminja ob spreminjanju standardnega odklona. V legendi imamo 2 različni ligi. Liga(1) predstavlja ligo, kjer se odigra eno dvokolo v sami ligi, medtem ko liga(2) označuje, da se odigrata dve dvokoli. To pomeni, da se v ligi(2) odvije dvakrat več iger kot v Ligi(1). Več tekem da favoritu lige večjo možnost, da osvoji prvo mesto. Jasno razvidno na grafu je, da Liga(2) dominira Ligo(1). Seveda bi liga(3) dala še boljše rezultate, vendar tega nismo prikazali, saj je več kot dve dvokoli na ligo redko srečamo v modernem športu.



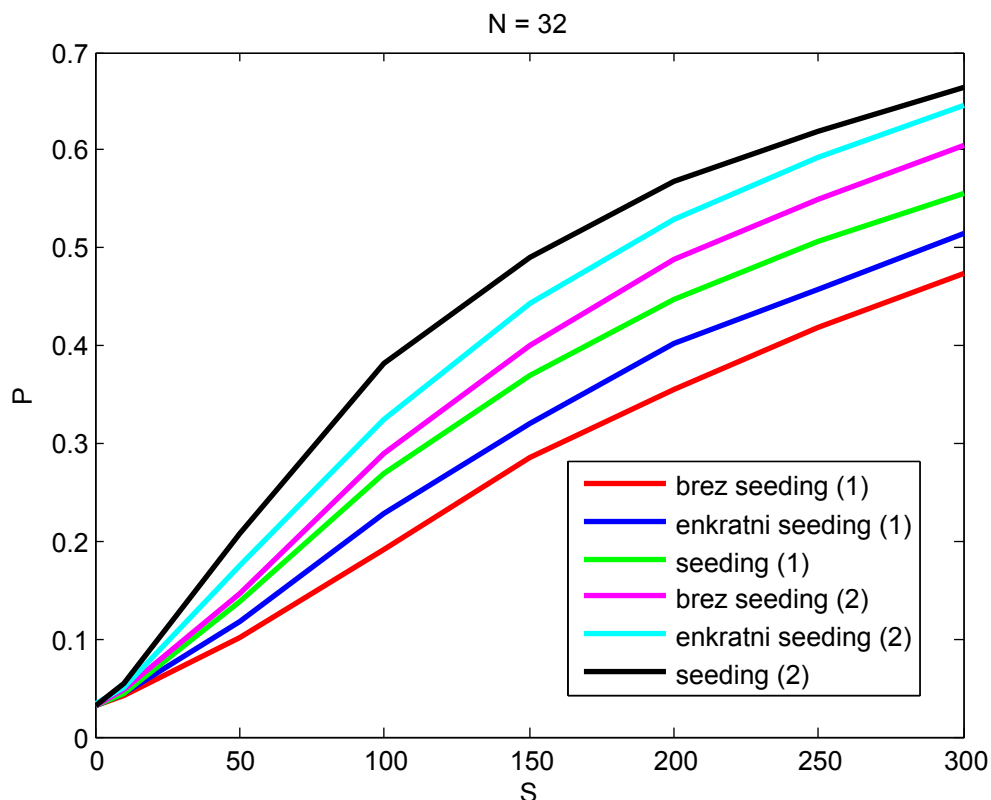
Slika 6.1: Verjetnosti zmage favoritov v ligi

6.2 Turnirji na izpadanje

V grafu na sliki 6.2 so predstavljene tri možnosti turnirja. Turnir lahko poteka brez seedinga (navaden žreb), lahko je turnir z začetnim seedingom (v legendi označen kot enkratni seeding) ali pa turnir z seedingom v vsakem koraku (v legendi poimenovan kot seeding). Števila v oklepaju predstavljajo število potrebnih zmag, da je nasprotnik premagan. Razlike so dokaj očitne. Turnir brez seedinga poda slabši rezultat kot turnir z enkratnim seedingom. Prav tako se da razbrati, da turnir s seedingom v vsakem koraku vrne najboljši rezultat.

Vpliv števila potrebnih zmag je tudi vidno prisoten. Več kot zahtevamo

zmag za napredovanje v naslednji krog, bolj verjetna postaja zmaga favorita. Ob boljšem pregledu smo opazili, da je število zmag bolj vpliven faktor na samo verjetnost, kot pa tip seedinga.

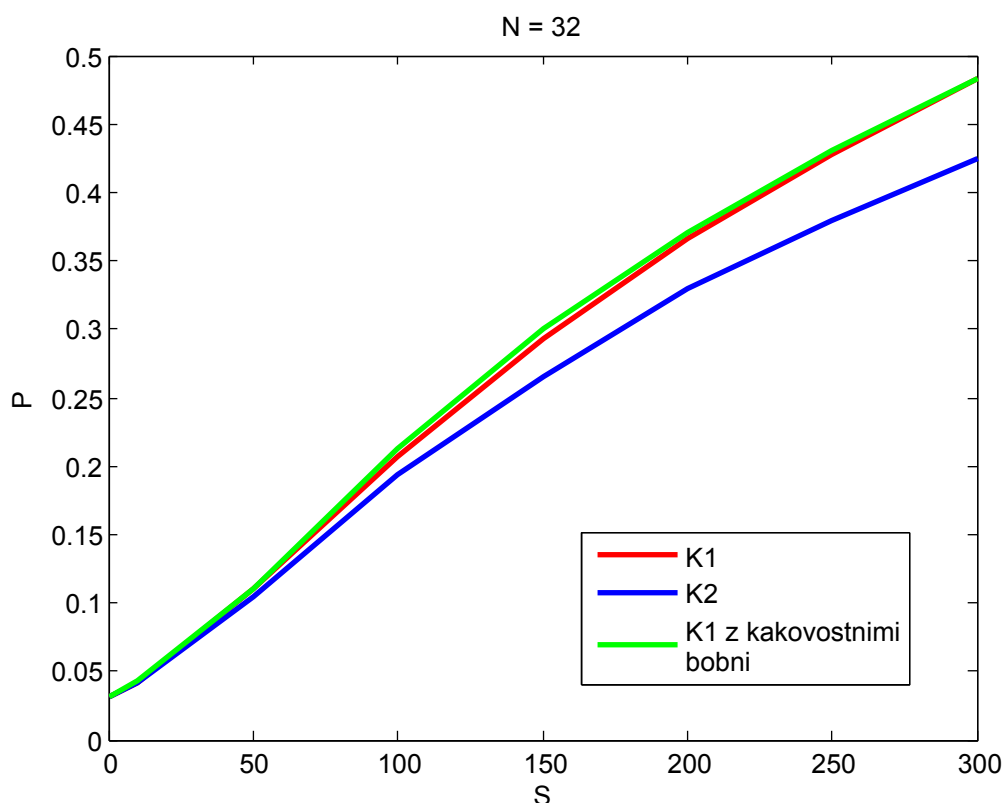


Slika 6.2: Verjetnosti zmage favoritov v turnirjih na izpadanje

6.3 Kombinirana tekmovanja

V grafu na sliki 6.3 smo ugotovili, da kombinirano tekmovanje K1 z kakovstnimi bobni rahlo izboljša tekmovanje K1. Za naš poskus K2 pa se izkaže da se ni obnesel, čeprav je dobro kazalo pri manjšem številu ekip, kot je razvidno v tabeli 6.2. Pri $N=8$ je K2 bolje deloval kot K1, nato pa že pri $N=16$

začne K1 prevzemati vodilno vlogo. Zanimivo pa je da kakovostni bobni v začetku ponekod celo škodujejo pri poskusu da bi dosegli boljši rezultat in celo znižajo verjetnost zmage favorita. Tu bomo samo omenili, da smo šli tudi do $N=128$ in da so kakovostni bobni že pri $N=16$ izboljšali tako K1 kot K2, vendar je treba poudariti da so izboljšave zelo skromne.



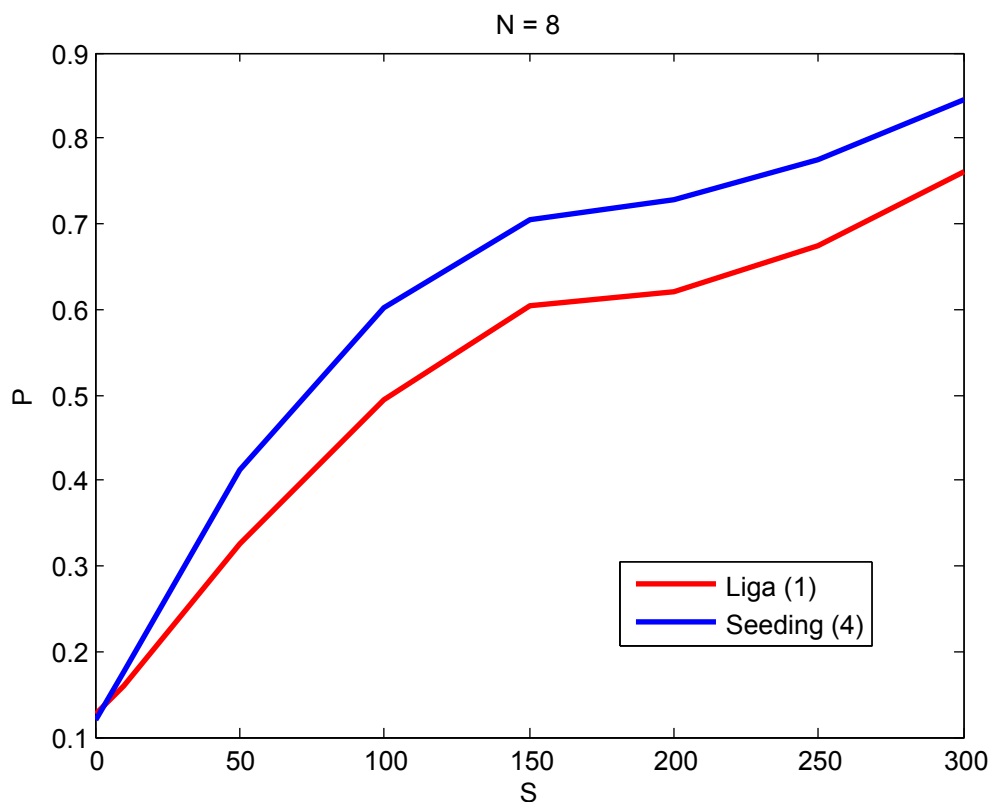
Slika 6.3: Verjetnosti zmage favoritov v kombiniranih tekmovanjih

Iz grafov 6.1, 6.2 in 6.3 razberemo, da se implementirana kombinirana tekmovanja ne morejo kosati z ostalimi. Tekmovanje liga je doseglo največje verjetnosti. Vendar pa je tudi razvidno, da se je najboljši med implementiranimi turnirji na izpadanje, seeding (2), najbolj približal Ligi(1). Po nekaj

¹kb je okrajšava za kakovostne bobne v tabeli 6.2.

tip tekmovanja (N)	S=10	S=50	S=100	S=150	S=200	S=250	S=300
K1(8)	0.146	0.234	0.351	0.444	0.499	0.574	0.608
K2(8)	0.146	0.247	0.365	0.462	0.515	0.582	0.617
K1(8) kb ¹	0.147	0.236	0.352	0.444	0.494	0.567	0.599
K2(8) kb	0.148	0.248	0.369	0.465	0.516	0.583	0.614

Tabela 6.2: Tabela verjetnosti zmage favorita v kombiniranih tekmovanjih.



Slika 6.4: Primerjava verjetnosti zmage favoritov

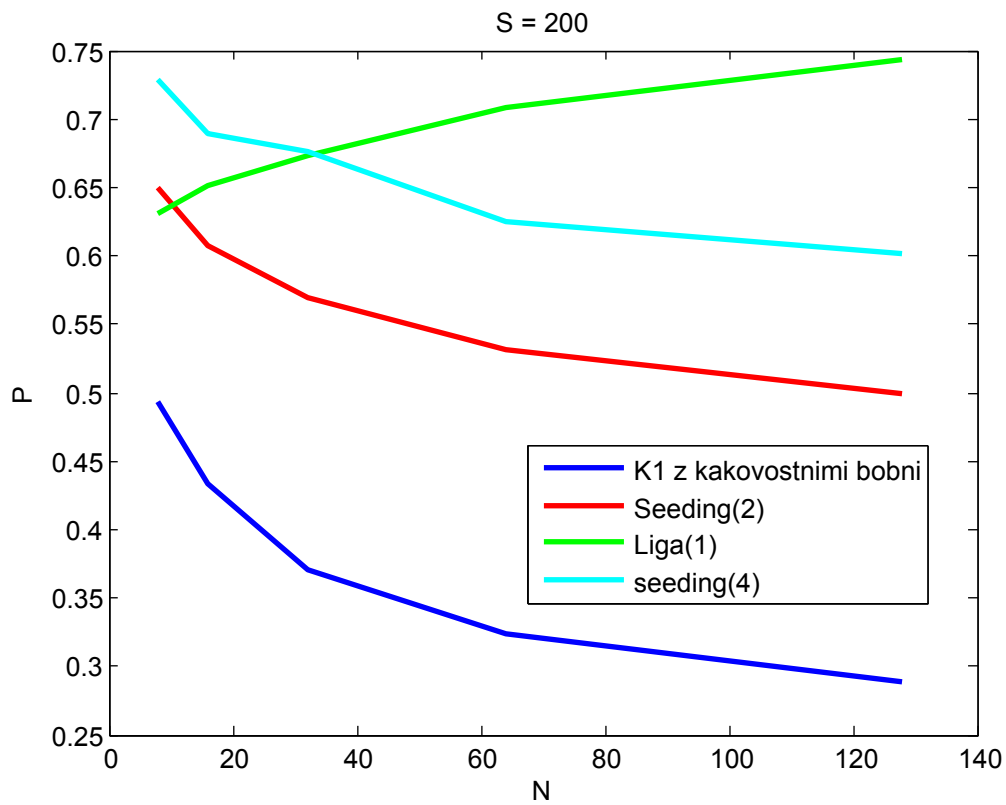
iskanja smo naleteli na ligo NHL, ki v zaključku svojega tekmovanja uporablja način, kot bi bil pri nas seeding(4). Liga NHL je trenutno razdeljena na

dve diviziji. V vsaki diviziji napreduje osem ekip. Torej po rednem delu NHL tekmovanja ali kot bi ga sedaj mi označili z liga, sledi še turnir na izpadanje, kjer se igra na 4 zmage za napredovanje. Kot naše zadnje dejanje poskusimo še primerjavo seeding(4) in lige(1).

In kot se da sedaj ugotoviti iz grafa na sliki 6.4 so pri NHL našli boljši način za določanje zmagovalca posamezne divizije, kot pa če bi kvalifikantje odigrali med seboj ligo. Vendar ko pogledamo malo globlje in poženemo še nekaj več iteracij je hitro jasno, da pri obravnani točki $N=64$ seeding(4) kmalu začne popuščati in liga spet prevzame vodilno vlogo, kar nam nakazuje tabela 6.3. Če pa si obrnemo pogled in fiksiramo S , kot je to vidno na sliki 6.5, pa se jasno vidi, da liga z povečevanjem parametra N edina pridobiva na verjetnosti zmage favorita, medtem ko vse ostale izgubljajo. Lahko bi probali povečati število potrebnih zmag. Sedaj že pri seeding(4) lahko en dvoboj traja tudi do 7 tekem. Kako monotono in nezanimivo bi bilo tekmovanje z dvoboji, ki trajajo po 30 ali pa še več tekem z istim tekmečem in pa za koliko dolgo bi to sploh bilo bolj učinkovito, preden spet liga prevzame svoj prestol.

	S=10	S=50	S=100	S=150	S=200	S=250	S=300
Liga(1)	0.048	0.271	0.453	0.550	0.625	0.694	0.722
Seeding(4)	0.069	0.410	0.597	0.674	0.705	0.731	0.737

Tabela 6.3: Primerjava verjetnosti dveh tekmovanj. Liga(1) že počasi dohiteva seeding(4) ($N = 64$).

Slika 6.5: Primerjava verjetnosti zmage favoritov v odvisnosti od N .

Poglavje 7

Sklepne ugotovitve in nadaljnje delo

Torej, iz poglavja 6 lahko določimo naš končni rezultat oziroma odgovorimo na vprašanje, kaj je najboljše za favorita. Najboljši način za določanje zmagovalca v tekmovanjih je liga. Poleg tega, da v grafih prejšnjega poglavja vidimo ligo kot zmagovalko, dajmo še malo premisliti zakaj je temu tako. V ligi ekipe namreč odigrajo več tekem in pa tudi proti vsem ekipam za razliko od preostalih tekmovanjih. To, da se pomerijo z vsemi pomeni, da si v večini primerov lig lahko privoščijo občasen poraz, kar pa je pri turnirjih na izpadanje lahko usodnega pomena (odvisno od števila zahtevanih zmag za napredovanje). To pa še ni vse. V ligi je zagotovljeno, da bo vsaka ekipa igrala z vsako, torej bo favorit tekmovanja zagotovo igral tudi z najslabšimi v ligi (lažje tekme). V turnirjih na izpadanje pa so pogosto najslabše ekipe izločene bolj na začetku in posledično napredujejo boljše. Torej bo favorit imel v povprečju manj lahkih nasprotnikov, medtem ko se v kombiniranih tekmovanjih najslabši tekmovalci bolj poredko uvrstijo na del tekmovanja z izpadanjem.

Ligi se najbolj približa turnir na izpadanje s seedingom v vsakem koraku, v kolikor dovolj povečamo število potrebnih zmag za napredovanje oziroma zmago, vendar pa na to močno vpliva število udeležencev. Več udeležencev

pomeni večji t.i. maraton za ekipo v ligi, kjer si lahko privošči poraz, v turnirjih na izpadanje pa pomeni več tekmecev, ki je je nujno treba premagati za končno zmago (zahteva po velikem številu zaporednih zmag narašča).

Kako težka so kombinirana tekmovanja, je nekaj kar nas je malce presenetilo. Če bi nekdo rekel, da ima neko moštvo kot favorit 35% možnosti za zmago v nekem sodobnem kombiniranem tekmovanju (recimo svetovno prvenstvo v nogometu), bi zagotovo po občutku pred diplomsko nalogo rekli, da je to število nekoliko nizko, vendar pa se izkaže, da je zelo realno. Prav tako sedaj lahko rečemo za isto tekmovanje bi bila ocena 40% že odlična, čeprav nam na prvi pogled to ni bilo očitno.

V nadaljevanju oziroma nadgrajevanju dela bi lahko upoštevali tudi prednost domačega igrišča in raziskali če in kako to vpliva na rezultate. Mogoče bi se splačalo pogledat v kakšni relaciji so števila odigranih tekem turnirjev in lige, kjer pri obeh tekmovanjih dosegamo podobne verjetnosti (primer bi lahko bil liga(1) in seeding(4) pri pogojih $N=64$ in $S=200 - 300$).

Literatura

- [1] (2014) NBA. Dostopno na:
http://en.wikipedia.org/wiki/National_Basketball_Association
- [2] (2014) NHL. Dostopno na:
http://en.wikipedia.org/wiki/National_Hockey_League
- [3] (2014) Športna Liga. Dostopno na
http://en.wikipedia.org/wiki/Sports_league
- [4] (2014) Pokalni sistem. Dostopno na:
http://sl.wikipedia.org/wiki/Turnirski_razporeditveni_sistemi
- [5] (2014) Seeding. Dostopno na:
[http://en.wikipedia.org/wiki/Seed_\(sports\)](http://en.wikipedia.org/wiki/Seed_(sports))
- [6] (2014) Svetovno prvenstvo v nogometu (FIFA world cup). Dostopno na:
http://en.wikipedia.org/wiki/FIFA_World_Cup
- [7] (2014) Kakovostni bobni (angleško pots). Dostopno na:
http://en.wikipedia.org/wiki/FIFA_World_Cup#Final_tournament
- [8] Rice John, "Mathematical statistics and data analysis. Cengage Learning, str. 31-36,2006.
- [9] (2014) Gaussova porazdelitev Dostopno na:
http://sl.wikipedia.org/wiki/Normalna_porazdelitev

- [10] (2014) ELO. Dostopno na:
http://en.wikipedia.org/wiki/Elo_rating_system
- [11] (2014) Arpad Elo. Več o njem na:
http://en.wikipedia.org/wiki/Arpad_Elo
- [12] (2014) Eclipse. Dostopno na:
<https://www.eclipse.org/downloads/>
- [13] (2014) Java. Dostopno na:
[http://en.wikipedia.org/wiki/Java_\(programming_language\)](http://en.wikipedia.org/wiki/Java_(programming_language))
- [14] (2014) Evropsko prvenstvo v košarki Dostopno na:
http://en.wikipedia.org/wiki/FIBA_EuroBasket#Competition_format