

TALLINNA TEHNIKAÜLIKOOL

Infotehnoloogia teaduskond

Informaatika instituut

IDU40LT

Andreas Lepamets 134317 IABB

**SPORDIENNUSTUSTEGA TEGELEVA
ETTEVÕTTE TÖÖPROTSESSI
AUTOMATISEERIMISE
VÕIMALUSED**

Bakalaureusetöö

Juhendaja: Enn Õunapuu

PhD

Dotsent

Tallinn

2016

Autorideklaratsioon

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

(kuupäev)

(allkiri)

Annotatsioon

Paljud väiksemad spordiennustustega tegelevad kihlveokontorid ostavad spordiennustuste koefitsiente ettevõtelt, mis tegelevad üldise kihlveokontorite turu jälgimisega ja turuga vastavuses olevate uute koefitsientide arvutamisega. Tihti on selliste koefitsiente müüvate ettevõtete tööprotsess manuaalne. See tähendab, et turu jälgimisega tegelevad inimesed ning arvutatud koefitsientide õigsus sõltub inimese kiirusest ja täpsusest. Käesolevas bakalaureusetöös on kirjeldatud erinevaid meetodeid, kuidas koefitsientide müügiga tegeleva ettevõtte tööprotsessi automatiseerida. Töös on selgitatud, kuidas koguda turul olevaid koefitsiente, kasutades selleks veebilehtedelt kraapimise ja serveritesse päringute tegemise meetodeid. Mõlema meetodi jaoks on töös tehtud ka lihtne prototüüp, mida on võimalik aluseks võtta täispika automatiseeritud tööprotsessi loomisel.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 29 leheküljel. Töö koosneb sissejuhatausest, kolmest sisupeatükist ja kokkuvõttest. Töös on kaks joonist ja üks tabel. Kokku viitab lõputöö kaheksale erinevale allikale.

Abstract

Several smaller sports-betting bookmakers buy their sports-betting odds from companies that keep an eye on the market odds and calculate new odds which are consistent to the market.. This type of companies often have a manual workflow, where the correctness of calculated odds usually depend on a human worker's speed and accuracy. In this thesis there are described different methods to automate this kind of workflow by writing a program which would ask the odds from the websites or straight from the servers of large bookmakers and then analyse them further. Simple prototypes are also implemented which could be used for fully developing the described automatic workflow.

The thesis is in Estonian and contains 29 pages of text. There is an introduction, three main chapters and a summary. The thesis contains two figures, one table and eight references.

Lühendite ja mõistete sõnastik

Panustaja	<i>Bettor</i> Isik, kes sõlmib kihlveokontoriga spordiennustus tehinguid.
HTML	<i>HyperText Markup Language</i> Veebilehtede kirjutamiseks mõeldud märkimiskeel.
Kihlveokontor	<i>Bookmaker</i> Juriidiline ettevõtte, mis omab õigust kihlvedude sõlmimiseks.
Koefitsient	<i>Odds</i> Peegeldab mingi sündmuse juhtumise tõenäosust.
Spordiennustamine	<i>Sports-betting</i> Spordivõistluse tulemuse peale raha paigutamist eesmärgiga saada tulu.
URL	<i>Uniform Resource Locator</i> Interneti aadress mis identifitseerib mingi kindla faili asukoha.

Jooniste nimekiri

Joonis 1. Konkreetne näide mängust SKA Rostovi ja Angusht Nazrani vahel. 17

Joonis 2. Ettevõtte X koefitsiendide haldamiseks ettenähtud programm. 20

Tabelite nimekiri

Tabel 1. Kihlveokontori Unibet lehel pakutavad kogu väravate arvu koefitsiendid 15

Sisukord

SISSEJUHATUS	9
1. PROBLEEMI DOMEEN	11
1.1 Spordiennustus.....	11
1.2 Kihlveokontor	12
1.3 Koefitsiendid	12
1.4 Reaalajas panustamine.....	15
2. ETTEVÕTTE X TÖÖPROTSESS	18
2.1 Reaalajas spordiennustuskoeffitsiente haldav töötaja.....	18
2.2 Koefitsientide haldamiseks mõeldud programm	19
2.3 Tööprotsessi puudused	20
3. AUTOMATISEERITUD TÖÖPROTSESSI KIRJELDUS	22
3.1 Kihlveokontoritelt koefitsientide saamine	22
3.1.1 Koefitsientide kraapimine veebilehtedelt	22
3.1.2 Koefitsientide küsimine serverist	23
3.2 Uute koefitsientide arvutamine.....	25
3.3 Automatiseeritud tööprotsessi puudujäägid.....	25
KOKKUVÕTE	26
SUMMARY	28

KASUTATUD ALLIKAD	29
LISA 1. SELENIUM	30
LISA 2. <i>GET</i> PÄRING	31

SISSEJUHATUS

Spordiennustuste tegemine on viimase 15 aastaga muutunud inimeste seas väga populaarseks (Keogh ja Rose, 2013). Tulu, mida panustaja spordiennustuste pealt võib teenida, sõltub panustatud summa suurusest ja kihlveokontori pakutud koefitsientidest. Koefitsient on murdarvuline väärtus, mis peegeldab kihlveokontori hinnangut harilikult spordivõistluse mingi lõpptulemuse tõenäosusele.

Suuremad kihlveokontorid arvutavad ise spordiennustuste koefitsiente. Paljudel väiksematel kihlveokontoritel selle töö tegemiseks ressursid puuduvad ning nad ostavad taolist teenust sisse ettevõtelt, kes sellega tegelevad. Käesolevas bakalaureusetöös vaadeldakse ettevõtet X, mis tegeleb reaaliajas käimasolevate spordivõistluste jaoks uute koefitsientide arvutamisega. Seda tehakse, lähtudes mitmete suurte kihlveokontorite koefitsientidest. Neid kasutades leitakse turu keskmised koefitsiendid. Selle abil tagatakse, et koefitsiendid, mida ettevõtte X kihlveokontoritele müüb, on turu keskmisega kooskõlas.

Ettevõtte X tööprotsess on enamasti manuaalne ning toetub suurel määral inimese kiirusele ja täpsusele. Et koefitsiendid muutuvad reaaliajas ja tihti suhteliselt kiiresti, siis võib inimesel tekkida andmete töötlemisel vigu või mahajäämusi. See omakorda võib pikas perspektiivis tuua ettevõttele kahju. Samuti peab ettevõtte hoidma tööl palju töötajaid ning sellest tulenevalt on tööjõukulud suured.

Eespool kirjeldatud probleemi aitaks kõrvaldada automatiseeritud süsteem, mis ei toetu nii palju töötaja tähelepanelikkusele, kuid on siiski töötaja poolt kontrollitav ja hallatav. Sellest lähtuvalt said käesoleva bakalaureusetöö ülesanded püstitatud järgnevalt:

1. Kirjeldada koefitsientide arvutamisega tegeleva ettevõtte X praegust manuaalset tööprotsessi.
2. Kirjeldada ettevõtte X potentsiaalset automatiseeritud tööprotsessi, mis koosneb kahest etapist: koefitsientide turu keskmise hindamiseks vajaminevate andmete kättesaamine ja saadud andmetest uute koefitsientide arvutamine.

Manuaalse tööprotsessi kirjeldamiseks tutvuti põhjalikult ettevõtte X töötaja tööülesannetega ning hetkel kasutusel oleva koefitsientide haldamiseks ettenähtud programmiga. Töös kirjeldatakse programmis olevaid parameetreid, mida töötaja oma tööülesannete täitmiseks kasutab. Samuti tutvustatakse manuaalse tööprotsessi põhilisi puuduseid.

Suurte kihlveokontorite andmebaasidest koefitsientide saamiseks kirjeldatakse kahte võimalust. Esmalt vaadeldakse koefitsientide kraapimist kihlveokontorite veebilehelt, kasutades veebilehtede testimise programmi Selenium. Selleks tutvuti veebilehtede ülesehitusega ja Seleniumi tööpõhimõtetega. Teisena kirjeldatakse koefitsientide küsimist otse kihlveokontorite serveritest. Antakse lühike ülevaade, kuidas serveri päringud (eelkõige GET) töötavad ning kuidas päringu vastustest on võimalik koefitsiente kätte saada. Mõlemast meetodist loodi töö raames ka lihtne prototüüp. Viimases osas kirjeldatakse, kuidas saadud koefitsientide põhjal uusi koefitsiente arvutada.

Töö alguses selgitatakse erinevaid spordiennustusega seonduvaid mõisteid ning seletatakse lahti, millised on kihlveokontorite ning panustajate roll antud valdkonnas. Teises peatükis selgitatakse ettevõtte X, mis tegeleb spordiennustuskoeffitsientidega, näitel, milline on manuaalselt hallatav süsteem ning tuuakse välja selle puudused. Kolmandas peatükis käsitletakse erinevaid võimalusi hetkel kasutatava manuaalse tööprotsessi automatiseerimiseks ning kirjeldatakse automatiseeritud tööprotsessi võimalikke puuduseid.

1. PROBLEEMI DOMEEN

Tänapäeval on spordivõistluste tulemuste peale kihla vedamine väga populaarne. Inimesed teevad seda nii meelelahutuse kui ka raha teenimise eesmärgil. Järgnevad alapeatükid tutvustavad lugejale spordiennustamise valdkonda, mida antud bakalaureusetöö käsitleb. Defineeritakse erinevaid spordiennustamisega seonduvaid mõisteid ning kirjeldatakse selle üldist olemust. Eraldi räägitakse pikemalt üha enam populaarsust koguvast reaalarajas panustamisest.

1.1 Spordiennustus

Spordiennustuse tegemine tähendab mingi kindla spordivõistluse tulemuse peale raha paigutamist eesmärgiga saada tulu. Käesolevas töös kasutatakse sünonüümidenäiteid spordiennustus ja kihlvedu. Sündmus on defineeritud kui mingi konkreetne spordivõistluse tulemus (näiteks jalgpallis koduväljakul mängiva meeskonna võit), mille peale saab kihla vedada. Spordiennustusel on kaks osapoolt: teenuse müüja ehk kihlveokontor ja teenuse tarbija ehk panustaja. Spordiennustus on äriühing nende osapoolte vahel. Selle tehingu käigus üritavad mõlemad oma tulu maksimeerida. Käesolevas töös vaatleme spordiennustust jalgpalli näitel.

Jalgpallile panustamine on alates 1998. aasta maailmameistrivõistlustest ja 2000. aasta Euroopa meistrivõistlustest muutunud väga populaarseks. Spordivõistlustele panustamise majandusharu on Darren Smalli arvates väärt 700 miljardit kuni 1 triljon dollarit aastas. Ligi 70% selles summast tuleneb jalgpallile panustamisest. (Koegh ja Rose, 2013.)

Jalgpallimatš võib lõppeda kolme üksteist välistava sündmusega: koduväljakul mängiva meeskonna võit, viik või võõrsil mängiva meeskonna võit. Lisaks nimetatud sündmustele pakuvad kihlveokontorid võimalust panustada ka väga paljudele muudele sündmustele, näiteks mängu käigus kokku löödud väravate arvule, nurgalöövide arvule,

esimesele väraalööjale, viimasele väraalööjale, või isegi sellele, kas teatud mängija hammustab teist mängijat või mitte (Rovell, 2014).

Kihlvedu on võimalik sõlmida kõigile eelpool mainitud mängu käigus või lõpuks toimunud sündmustele. Spordiennustust nimetatakse tulusaks, kui toimub just see sündmus, millele raha panustati. Tulu suuruse määrab rahasumma, mille panustaja panustas ning kihlveokontori määratud koefitsient, mis on seotud kihlveokontori ennustatud sündmuse juhtumise tõenäosusega.

1.2 Kihlveokontor

Üldiselt võib kihlveokontoriks nimetada kõiki inimesi või organisatsioone, kes pakuvad spordisündmustele koefitsiente ning võimalust sündmuse erinevate tulemuste peale panustada. Käesoleva töö kontekstis tähendab kihlveokontor aga juriidilist kihlvedude sõlmimise õigust omavat ettevõtet. Kihlveokontor pakub koefitsiente oma kasu saamise eesmärgil ning selle saavutamiseks seab need vastavalt. Turuks nimetatakse kõiki kihlveokontoreid, kes omavad eelpool mainitud õigusi ning turu keskmisteks koefitsientideks kõikide kihlveokontorite poolt pakutava mingi kindla sündmuse keskmist koefitsienti. Näiteks kui turg koosneks kõigest kahest kihlveokontorist, kus üks pakub sündmuse A koefitsiendiks 1,80 ja teine pakub sündmuse A koefitsiendiks 1,90, oleks turu keskmine koefitsient sündmusele A 1,85. Mõned populaarseimad kihlveokontorid Euroopas on Bet365, Williamhill ja Bwin. Aasia populaarseimad veebipõhised kihlveokontorid on aga 188bet ja SBOBET.

1.3 Koefitsiendid

Koefitsient peegeldab seda, kui suur on kihlveokontori hinnangul mingi sündmuse juhtumise tõenäosus. Näiteks kahe võrdse tõenäosusega juhtuva sündmuse koefitsiendid on ligikaudu võrdsed.

Spordiennustustes kasutatakse kahte tüüpi koefitsiente. Inglise tüüpi koefitsiente esitatakse hariliku murruna. See näitab sündmuse toimumisel saadavat kasu, arvestades panuse suurust. Näiteks inglise tüüpi koefitsient 5/1 näitab, et ühe ühiku panustamisel

on oodatav kasum viis ühikut. Seega kokku saab panustaja raha tagasi kuue ühiku ulatuses. Euroopa tüüpi koefitsiente esitatakse kümnendmurruna. Sel juhul saadakse lõplik kihlveokontori poolt väljamakstav summa panuse korrutamisel koefitsiendiga. Näiteks inglise tüüpi koefitsiendile 5/1 vastab euroopa tüüpi koefitsient 6. Ühe ühiku panustamisel saab panustaja tagasi $1 \cdot 6 = 6$ ühikut raha. Euroopa tüüpi koefitsiendi pöördväärtus võrdub sündmuse toimumise tõenäosusega. Koefitsiendi väärtuse 6 korral on sündmuse toimumise tõenäosus seega $1/6 = 16,67\%$. Käesolevas töös kasutatakse edaspidi ainult euroopa tüüpi koefitsiente. (Buchdal, 2003.)

Et kihlveokontorid on äriettevõtted, mis on ka ise huvitatud kasumi saamisest, siis pakuvad nad kasutajale nõ ebaausaid koefitsiente, mis tagavad kihlveokontorile pikas perspektiivis võidu, olenemata sellest, millise seisuga spordisündmused lõppevad (Buchdal, 2003). Üksteist välistavate sündmuste koefitsiente nimetatakse ausaks, kui nende pöördväärtuste (sündmuste tõenäosuste) summa võrdub ühega. Vaatame näiteks jalgpallimatši, kus tõenäosus, et lüüakse üle või alla 2,5 värava on võrdne (kihlveokontorid kasutavad murdarvulisi väravate arve, sest võimatu on lüüa näiteks 2,5 väravat ja see välistab olukorra, kus teatud tingimustel kumbki sündmus ei toimuks). Sellises olukorras oleks mõlema sündmuse koefitsient 2 (sündmuste juhtumise tõenäosus oleks 1/2) ehk panustaja saaks võidu korral oma panustatud rahasumma kahekordselt tagasi. Tegemist on ausate koefitsientidega, sest $1/2 + 1/2 = 1$. Ausad oleksid koefitsiendid ka siis, kui alla 2,5 värava ennustatud koefitsiendiks oleks 5 ja üle 2,5 värava koefitsiendiks 1,25, sest $1/5 + 1/1,25 = 1$. Sageli pakuvad kihlveokontorid aga kahe võrdvõimaliku üksteist välistava sündmuse koefitsientideks 1,83-1,95 vahemikku jäävaid suuruseid. Sellisel juhul on nende sündmuste summaarne tõenäosus ühest suurem. Tegemist on ebaausate koefitsientidega, mis tagavad selle, et kihlveokontori väljamakstav kogusumma on alati tema tulust väiksem.

Konkreetne näide koefitsientidest on toodud tabelis 1. Seal on näidatud koefitsiendid sündmustele, et matši jooksul lüüakse üle või alla 0,5, 1,5, 2,5 jne värava. Näiteks koefitsient sellele, et lüüakse rohkem kui 3,5 väravat on 5,2. Samuti on tabelist 1 näha, et kihlveokontor kasutab ebaausaid koefitsiente. Näiteks see, et mängu jooksul lüüakse rohkem kui 0,5 väravat ja vähem kui 0,5 väravat on teineteist välistavad sündmused, millest üks kindlasti toimub. Nende koefitsientide pöördväärtuste (sündmuste toimumistõenäosuste) summa on aga ühest suurem: $1/1,12 + 1/7,0 = 1,036$.

Kihlveokontorid arvutavad sageli koefitsiendid ümber, kui selguvad uued ühe või teise meeskonna võitu ennustavad asjaolud. Näiteks kui Portugali koondis mängib Eesti koondisega ning paar tundi enne mängu algust selgub, et Portugali koondise staarmängija on saanud vigastada ning kohtumist kaasa teha ei saa, muutub Portugali koondise võidu tõenäosus väiksemaks ning Eesti koondise oma vastavalt suuremaks ning see kajastub ka kihlveokontorite poolt pakutavates koefitsientides.

Samuti sõltuvad ühe konkreetse kihlveokontori koefitsiendid mingil määral teiste kihlveokontorite omadest, sest nad ei saa lubada arbitraaži. Arbitraažiks nimetatakse olukorda, kus panustajal on võimalik teenida kasumit olenemata sellest, milline sündmus üksteist välistavate sündmuste hulgast toimub. Vaatleme näitena jalgpallimatši, kus kihlveokontor A pakub koefitsienti 2,1 sündmusele, et mängu jooksul lüüakse kokku enam kui 2,5 väravat ja koefitsienti 1,67 sündmusele, et lüüakse vähem kui 2,5 väravat. Kihlveokontor B pakub aga vastupidiseid koefitsiente ehk üle 2,5 värava koefitsient on 1,67 ja alla 2,5 värava koefitsient on 2,1. Sellises olukorras on panustajal võimalik näiteks panustada kihlveokontoris A kümme eurot sündmusele, et lüüakse üle 2,5 värava (koefitsiendiga 2,1) ning kihlveokontoris B kümme eurot sündmusele, et lüüakse alla 2,5 värava (samuti koefitsiendiga 2,1). Panustaja panustab kokku 20 eurot ning kuna nendest sündmustest üks kindlasti toimub, saab ta ühelt või teiselt kihlveokontorilt tagasi 21 eurot. Seega tagab panustaja omale antud koefitsientide korral 5% kasumi kogu panustatud summast. Arbitraaži olukordi esineb üsna tihti, kuid need eemaldatakse kiirelt ning see on ka üks peamistest põhjustest, miks kihlveokontorid oma koefitsiente üldise turu koefitsientidega kooskõlas hoiavad. (Betting Arbitrage).

Suured kihlveokontorid kasutavad võimalikult täpsete sündmuste toimumise tõenäosuste ennustamiseks keerukaid algoritme. Väiksemad kihlveokontorid ostavad sageli aga teenust ettevõtetelt, kes tegelevad kogu turu koefitsientide jälgimisega ja nende põhjal uute arvutatud koefitsientide müümisega.

Tabel 1. Kihlveokontori Unibet lehel pakutavad kogu väravate arvu koefitsiendid Euroopa Meistriliiga finaalile 2016 Real Madrid vs Atletico Madrid (9.05.2016).

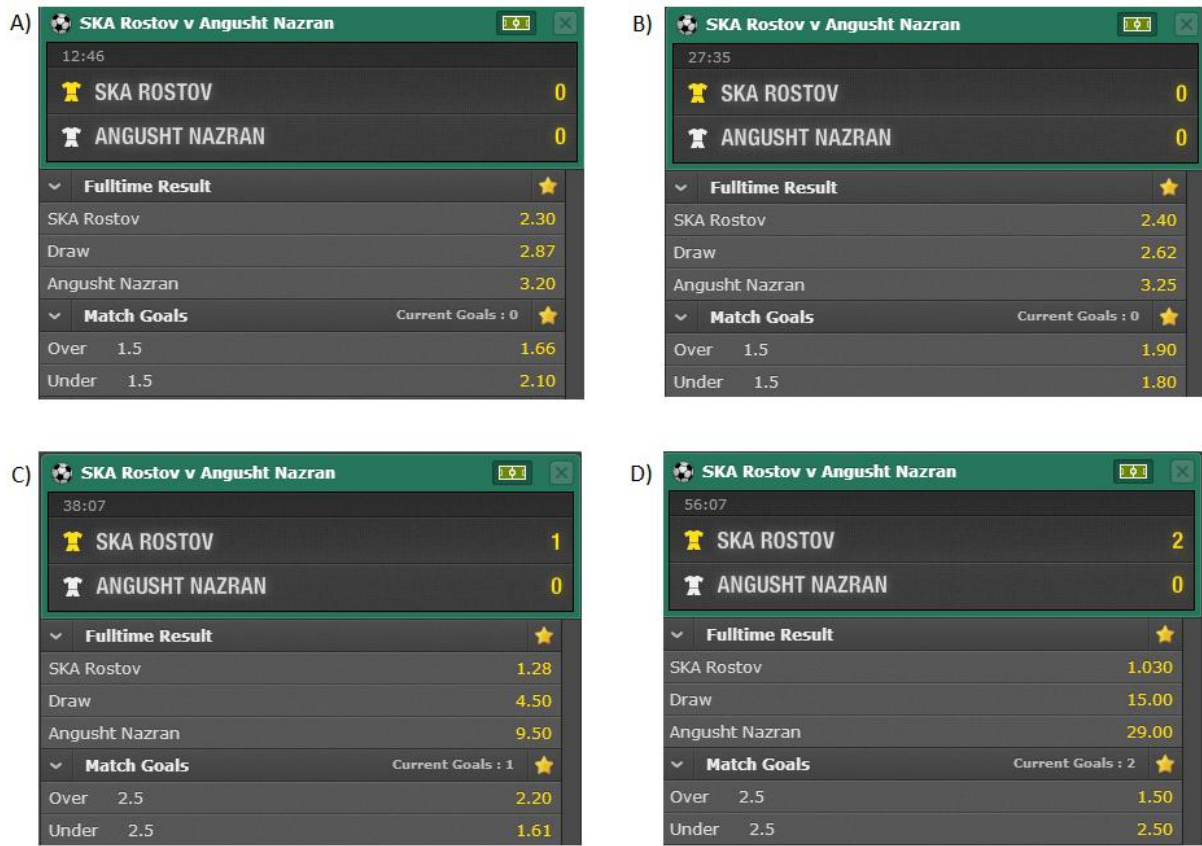
Väravad kokku			
Üle 0,5	1,12	Alla 0,5	7,00
Üle 1,5	1,51	Alla 1,5	2,65
Üle 2,5	2,55	Alla 2,5	1,54
Üle 3,5	5,20	Alla 3,5	1,18
Üle 4,5	12,00	Alla 4,5	1,06
Üle 5,5	32,00	Alla 5,5	1,01

1.4 Reaalajas panustamine

Kihlveokontorite veebilehtedel on võimalik panustada spordisündmustele enne selle algust, kuid üha enam on kihlveokontorite ja nende klientide seas populaarsust kogumas „live betting“ ehk sündmustele reaalajas panustamine. See tähendab, et panustamine toimub ajal, mil mäng juba käib. Selline panustamise viis on klientide seas populaarne enamasti seetõttu, et võimaldab panustajal enne panuse tegemist jälgida kohtumise mängu käiku ja muid detaile, mis võiks tal aidata mängu lõpptulemust ennustada. Siinkohal on oluline mainida, et iga panustaja potentsiaalselt saadav tulu sõltub just sellest koefitsiendist, mis kehtib sel momendil, kui panus tehti ja ei muutu, kuigi koefitsiendid mängu jooksul võivad muutuda. (Anderson *et al.*, 2012).

Vaatleme näitena jalgpallimatši kahe meeskonna vahel, kus enne mängu on mõlema meeskonna võidukoefitsiendid võrdsed (näiteks kummagi meeskonna võidu koefitsient

on 2,6 ja viigi koefitsient on 2,8). See tähendab, et kihlvekontor peab mõlema meeskonna võidu tõenäosust enam-vähem võrdseks. Kui mängu jooksul peaks kumbki meeskond värava lööma ning kohtumist juhtima minema, muutuvad ka reaalses panustamise koefitsiendid, sest juhtima läinud meeskonda peetakse sellest hetkest mängu favoriidiks. Selle meeskonna koefitsient muutub väiksemaks ja vastasmeeskonna oma vastupidiselt suuremaks. Selline süsteem annab panustajale, kes usub, et kaotusseisu jäänud meeskond suudab siiski kohtumise võita, suurema koefitsiendi ja potentsiaalselt suurema võidu. Kirjeldatud protsess ei toimu ainult mängu võitja ennustamise koefitsientidega, vaid ka eeldatavate väravate koguarvu koefitsientidega. Näiteks, kui enne mängu pakub kihlveokontor, et jalgpallimatšil lüüakse kogusummas üle 2,5 värava koefitsiendiga 1,83 ning mängu algusjärgus väravat ei lööda, hakkab see koefitsient aeglaselt tõusma. Kiire värava korral mängu alguses oleks antud sündmuse koefitsiendi muutus aga vastupidine. Reaalne näide koefitsientide muutumisest mängu käigus on toodud joonisel 1.



Joonis 1. Konkreetne näide mängust SKA Rostov ja Angusht Nazrani vahel.

Kihlveokontor Bet365 pakkus mängu 13. minutil järgnevaid koefitsiente, kus „Draw“ tähendab viiki, „Match Goals Over 1,5“ üle 1,5 kogu väravate summat ja „Match Goals Under 1,5“ alla 1,5 kogu väravate summat. Mängu seis oli 0:0 ning SKA Rostovi võidu koefitsient oli 2,30, viigi koefitsient 2,87 ja Angusht Nazrani võidu koefitsient 3,20. Koefitsient sündmusele, et mängus lüüakse üle 1,5 värava oli 1,66 ning alla 1,5 värava 2,10 (vt joonis A). Samas mängus 15. minutit hiljem polnud endiselt löödud ühtegi väravat ning selle tõttu oli kihlveokontor koefitsiente muutnud. Viigi koefitsient oli muudetud väiksemaks, sest mängu seis oli viigis. Samuti oli kihlveokontor langetanud koefitsiente sündmusele, et mängus lüüakse alla 1,5 värava (vt joonis B). Mängu 37. minutil lõi SKA Rostov värava ning selle järel muutis kihlveokontor nende võidu koefitsienti väiksemaks ning vastaste oma vastavalt suuremaks. Samuti muutis kihlveokontor ka kogu väravate summa koefitsiente (vt joonis C). 56. minutil lõi SKA Rostov ka teise värava. Kui mängu 13. minutil pidas kihlveokontor SKA Rostovi võidu tõenäosust umbes 43%, siis pärast 37. minutil löödud väravat oli nende võidu tõenäosus ligi 78% ning pärast 56. minutil löödud teist väravat oli nende võidu tõenäosus juba lausa 97% (vt joonis D).

2. ETTEVÖTTE X TÖÖPROTSESS

Olgu meil ettevõtte X, mis tegeleb koefitsientide müümisega väiksematele kihlveokontoritele ning selleks on palgatud vastavad töötajad, kes jälgivad turu koefitsiente. Järgnev peatükk teeb ülevaate ettevõtte X spordiennustuskoefitsientide haldava töötaja tööülesannetest. Samuti tuuakse siin välja hetkel järgitava manuaalse süsteemi kitsaskohad ja puudujäägid.

2.1 Reaalajas spordiennustuskoefitsientide haldav töötaja

Reaalajas spordiennustuskoefitsientide haldava töötaja ehk halduri tööülesandeks on tagada ettevõtte X haldusprogrammis olevate ja turu keskmiste koefitsientide vastavus mingi kindla jalgpallimatši kõikidel pakutavatel sündmustel.

Halduri töö seisneb turul olevate koefitsientide jälgimises ja nendest lähtuvalt uute koefitsientide arvutamises. Et tervet turgu on võimatu jälgida, kasutab haldur oma töös suuremate veebipõhiste kihlveokontorite koefitsiente. Sobivateks kihlveokontoriteks on näiteks Sbobet, 188bet, Bet365, WilliamHill ja teised.

Sageli haldab ettevõtte X koefitsientide haldur korraga ligi kuut mängu, jälgides nende tarbeks kokku ligi 20 veebilehte. Haldur peab eri veebilehtedel jälgima samade sündmuste koefitsiente. Enamasti on jälgitavateks sündmusteks mängu lõpptulemuse ja kokku löödud väravate arv. Näiteks kui kihlveokontor A, mida haldur jälgib, pakub mingi sündmuse koefitsiendiks 1,60 ja kihlveokontor B selle sama sündmuse koefitsiendiks 1,70, peab haldur tagama ettevõtte X koefitsientide haldamiseks mõeldud programmis selle sama sündmuse juhtumise koefitsiendiks ligikaudu turul olevate koefitsientide keskmise ehk 1,65.

2.2 Koefitsientide haldamiseks mõeldud programm

Ettevõttel X on olemas programm, mis hõlbustab halduri tööd. Programm ise arvutab koefitsiendid, aga halduri ülesandeks on jooksvalt programmi sisendit selliselt muuta, et arvutatud koefitsiendid vastaks enam-vähem turu keskmisele. Praegune programm lihtsustab mõnevõrra koefitsientide halduri tööd, kasutades algoritmi, mis muudab ka iseseisvalt koefitsiente mängu kestvusest või löödud väravatest sõltuvalt. Näiteks kui 10 minutit pärast mängu algust pole mängus väravaid löödud ning haldur pole programmi parameetreid muutnud, on koefitsiendid programmis siiski mõnevõrra muutunud, sest tõenäosus, et kumbki meeskond lööb värava, muutub ajaga aina väiksemaks. Samuti muutuvad koefitsiendid automaatselt kui üks meeskondadest lööb värava.

Ettevõtte koefitsientide haldamiseks mõeldud programmis on halduril võimalik muuta koefitsiente kahe parameetri ehk muutuja abil. Esimene parameetritest on „*Total Goals*“, mille abil on võimalik muuta koefitsiente mängus kokku löödavate väravate arvule. Mida suurem on tõenäosus, et mängus lüüakse palju väravaid, seda suuremaks peab haldur selle muutuja muutma. Teiseks halduri poolt muudetavaks parameetriks on „*Goal Supremacy*“, mis määrab koduväljakul mängiva meeskonna eeldatavat paremust. See muutuja võib olla ka negatiivne ning sel juhul ennustatakse võitjaks võõrsil mängivat meeskonda. Mida suurem on selle muutuja absoluutväärtus, seda väiksem on favoriidi ja suurem tema vastase võidu koefitsient.



Joonis 2. Ettevõtte X koefitsientide haldamiseks ettenähtud programm.

Joonise ülemisel paneelil on kujutatud jalgpallimatši kahe meeskonna, Santos FC ja Audax EC, mängu statistikat (mänguseis, nurgalöökid, kollaste ja punaste kaartide arv). Keskmisel paneelil on kujutatud kahte parameetrit „Total Goals“ ja „Goal Supremacy“, mida koefitsientide haldur muudab, et tagada alumisel paneelil kujutatud sündmuste koefitsientide vastavus turu keskmisega kindlal ajahektel. Alumisel paneelil on kujutatud mängu kõikide sündmuste koefitsiendid. Näiteks Santos FC võidu koefitsient on vaadeldaval hetkel 1,50.

2.3 Tööprotsessi puudused

Kuigi praegune protsess, kus haldur jälgib turul olevaid koefitsiente ning kasutab uute koefitsientide arvutamiseks selleks ettenähtud programmi, on töövõimeline, on sellel mitmeid puudujääke. Ettevõtte jaoks on suurimaks miinuseks suur tööjõu kulu. Kevadel

ja sügisel, kui enamustes maailma jalgpalli liigades hooajad käivad, on nädalavahetusesti korraga käimas väga suurel hulgal jalgpallimatše, mis vajavad haldamist. Näiteks on hooaja kõige tihedamal ajal, ühel laupäeval päeval kell 17:00, alanud korraga ligi 150 jalgpallimatši. Nende kõikide haldamine vajab aga paljusid haldureid, kellele on vaja maksta ka palka.

Samuti on probleemiks ka töötajate väsimus. Sageli, kui tööpäevad on pikad, langeb ka töötajate keskendumisvõime ning täpsus. See toob omakorda kaasa koefitsientide muutuste mittetähelepanemist ning vead võivad kergelt tekkida, sest koefitsiendid muutuvad tihti. Lisaks sellele ei suuda haldur ka nii kiiresti ettevõtte X programmis olevaid koefitsiente muuta, mis võib olla ettevõttele kahjulik, sest kliendid ootavad ettevõttelt täpsust just turu keskmiste koefitsientide näol.

3. AUTOMATISEERITUD TÖÖPROTSESSI KIRJELDUS

Antud töö eesmärgiks on uurida võimalusi eelpool kirjeldatud tööprotsessi automatiseerimiseks, mis aitaks vältida manuaalsest tööprotsessist tulenevaid probleeme. Automatiseeritud süsteem vajab edukaks toimimiseks vähem tööjõudu ning suudab vältida vigu, mis tekivad halduri ebatäpsusest.

Antud automatiseeritud süsteemi on võimalik jagada kaheks erinevaks etapiks. Esimene etapp hõlmab endas andmete ehk koefitsientide kogumist ning teine etapp nende töötlemist ja edasist analüüsi.

3.1 Kihlveokontoritelt koefitsientide saamine

Selles peatükis käsitletakse kahte erinevat meetodit kihlveokontoritelt koefitsientide saamiseks. Esimene meetod on kraapimine veebilehtedelt, kasutades selleks Seleniumit ning teine meetod on päringute tegemine kihlveokontorite serverist.

3.1.1 Koefitsientide kraapimine veebilehtedelt

Veebilehelt kraapimine tähendab andmete eraldamist veebilehe lähtekoodist. Käesoleva peatüki eesmärgiks on näite põhjal kirjeldada, kuidas on võimalik eraldada kihlveokontori veebilehe lähtekoodist koefitsiendid ning nendega seotud andmed.

Kihlveokontorite lehtedelt lähtekoodi kraapimiseks on võimalik kasutada programmi Selenium. Seleniumi põhiliseks kasutusala on lihtsustada veebilehete testimist, võimaldades erinevaid lahendusi ja tööriistu selle automatiseerimiseks. Sellega on võimalik avada veebilehitseja ning anda käsklusi erinevate ülesannete täitmiseks antud veebilehel. Kuna Selenium suudab lihtsustada veebilehtedelt andmete kogumist, kasutati antud töös just selle programmi võimalusi. (SeleniumHQ).

Bakalaureusetöös kasutati Seleniumit programmeerimiskeeles Python arvutis, mille operatsioonisüsteemiks oli Windows 7. Selenium paigaldati, kasutades käsuri (*cmd.exe*) ning selleks kasutati Pythoni pakettide paigaldamiseks mõeldud programmi *pip*. Seleniumi paigaldamise juhend asub veebilehel (Selenium Master).

Kihlveokontori veebilehelt koefitsientide kraapimise võimalikkuse näitamiseks loodi lihtne prototüüp, mis kraabib andmeid SBOBET-i veebilehelt www.sbobet.com¹. Programm vajab sisendiks mõne SBOBET-i lehel käimasoleva jalgpallimatši reaalajas panustamise veebilehe aadressi (URL-i). Veebilehe HTML lähtekood on üles ehitatud hierarhiliselt. See koosneb märgenditest (ingl. k *tag*), atribuutidest (ingl. k *attribute*) ja stiililehtedest (ingl. k *stylesheet*). Igal märgendil võib olla klass ja/või unikaalne ID väärtus, mille abil saab seda lähtekoodist üles leida. Näiteks otsitakse antud prototüübis veebilehe lähtekoodist üles mängu tulemuse koefitsientide tabel (märgend `<table>`), kasutades selle ID-d ning itereeritakse üle antud tabeli ridade (märgendid `<tr>`), et kätte saada selles tabelis olevad koefitsiendid. Prototüüp otsib veebilehelt mängu tulemuste koefitsiente ning prindib need standardväljundisse. Kommenteeritud programmi kood on lisas 1.

Antud lahendus ei pruugi olla kõige optimaalsem, sest veebilehtedelt kraapimine on suhteliselt aeglane protsess. Seda eelkõige seetõttu, et veebilehed sisaldavad ka palju muid andmeid, mille laadimine võtab kaua aega, aga mida tegelikult antud eesmärgi täitmiseks vaja ei ole. Prototüübi käivitamisel läheb esimeste koefitsientide kätte saamiseks ligi 20 sekundid. Selle ajaga avab ta Mozilla Firefox veebibrauseri ning laeb etteantud veebilehel olevaid andmeid. Kiirem lahendus oleks küsida koefitsiente otse kihlveokontori serverist.

3.1.2 Koefitsientide küsimine serverist

Kihlveokontorite andmebaasides muutuvad koefitsiendid pidevalt. Selle kajastamiseks kihlveokontorite veebilehtedel tehakse nende serveritesse päringuid, mis küsivad

¹ Sbobet.com kihlveokontori veebilehel on Eesti Vabariigis kihlvedude sõlmimine keelatud (hasartmänguseadus §56) ning sellest tulenevalt on veebilehele juurdepääs Maksu- ja Tolliameti poolt piiratud. Veebilehele on siiski võimalik ligipääseda, muutes manuaalselt oma arvutis kasutatavaks DNS serveriks näiteks Google-i DNS serveriks. Käesoleva töö raames ei ole Eesti Vabariigi seadusi rikutud, sest töö autor ei ole nimetatud kihlveokontoris kihlvedusid sõlminud.

serverilt koefitsiente mingi kindla mängu kohta ja mingil kindlal ajahetkel. Server tagastab päringu peale vastuse uute koefitsientidega ning need kuvatakse veebilehtedel.

Enamlevinudateks päringu tegemise meetoditeks on GET ja POST (Korpela 2003). Kihlveokontorite veebilehed kasutavad enamasti GET meetodit. See on POST meetodist vähem turvaline. Kõik parameetrite nimed ja väärtused kajastuvad päringu URL-s. Kihlveokontorite päringud sisaldavad harilikult parameetritena konkreetse mängu ID-d ja käesolevat ajahetke. Lisaks võivad nad sisaldada ka muid parameetreid. Server tagastab nendele parameetritele vastavad koefitsientide väärtused.

Vaatleme kihlveokontorit SBOBET ja mängu Czech Republic U23 ja Mali U23 vahel. Kihlveokontori veebileht saadab serverile järgmise päringu:

```
https://www.sbobet.com/en/data/event?ts=a9199619&tk=2873,2873,25,1,233,19891,1900121,20160518,1,4,0,4&vs=0
```

Päring algab serveri aadressiga ning pärast küsimärki tulevad &-märgiga eraldatud GET meetodi parameetrid (antud juhul „ts“, „tk“ ja „vs“) ja nende väärtused. Server saadab päringule vastava vastuse:

```
$M('tb').onDT('201605182350');$P.onUpdate('od',[2875,0],[[1],[1,19891,[1900121,'Czech Republic U23','Mali U23','1.009-E126016051801',10,'05/18/2016 23:00','',0,0,0,10,{}],[1,5,7,45,0,0,{},'HT',{1:40,2:40,3:15,4:15},0],[[5,5,[]],[23331562,[1,0,1,1,2000,0.00,1639840],[1.62,2.42]],[23331563,[1,0,1,1,3000,0.25,1639840],[2.17,1.77]],[23331565,[3,0,3,1,3000,1.00,1639840],[1.67,2.26]],[23331566,[3,0,3,1,2000,1.25,1639840],[2.12,1.77]],[23331568,[5,0,5,1,2000,0.00,1639840],[2.62,2.23,4.00]]],1],[[],[]],,,1]);$P.onUpdate('ls',[0,0,0]);
```

Päringu vastus sisaldab jalgpallimatši osapooli ja mängu toimumise aega. Põhiline osa vastusest koosneb erinevate sündmuste koefitsientidest. Näiteks punase värviga on esile tõstetud mängu kõigi võimalike lõpptulemuste koefitsiendid ning sinise värviga on märgistatud antud sündmuste identifikaator konkreetse mängu korral. Kasutades seda identifikaatorit, on kerge vajaminevaid koefitsiente serveri vastusest üles leida.

Sarnaselt veebilehele on päringuid võimalik teostada ka koefitsientide väärtuste kätte saamiseks suvaliselt IP-aadressilt ja muudeks eesmärkideks. Selleks võib kasutada näiteks Pythoni paketti *urllib2*. Lühike näide on toodud lisa 2. Selline meetod on kiirem kui veebilehelt kraapimine. Ühe päringu saatmine ja vastuse saamine võtab kokku umbes 350 ms.

3.2 Uute koefitsientide arvutamine

Kui koefitsiendid on kihlveokontorite andmebaasidest kokku kogutud, tuleb neid analüüsida. Selleks tuleks luua programm, mis võrdleks erinevate kihlveokontorite samade sündmuste tõenäosuseid ja koefitsiente ning arvutaks nende keskmise. Samuti oleks selle programmi abil võimalik ka avastada olukordi, kus kihlveokontorite pakutavad koefitsiendid on omavahel abitraažis. Sellisel juhul tuleks ajutiselt koefitsientide edastamine klientidele peatada.

Kirjeldatud programmi saaks ühildada ettevõttes X olemasoleva programmiga. Üheks võimaluseks oleks tööprotsess, kus uus programm hindab turu keskmised koefitsiendid ning arvutab nende põhjal ettevõtte X programmi jaoks „*Total Goals*“ ja „*Goal Supremacy*“ parameetrite väärtused. Sellisel juhul saaks haldur neid samu parameetreid kasutades koefitsiente endiselt vastavalt klientide soovidele vajadusel väiksemaks või suuremaks muuta. Seega on inimfaktor endiselt oluline ka automatiseeritud süsteemis, et programmi tööd jälgida, kuid inimese roll on palju väiksem. Tegelikult on võimalik kasutada masinõppe meetodeid, et õpetada programm ise koefitsiente muutma vastavalt mängu statistikale või sellele, kuidas kliendid eelistavad.

3.3 Automatiseeritud tööprotsessi puudujäägid

Eespool kirjeldatud automatiseeritud tööprotsess omab ka negatiivseid külgi. Kihlveokontorid on enamasti suhteliselt tundlikud kõige suhtes, mis nende veebilehtedel tehakse või milleks nende andmeid kasutatakse. Kuigi nimetatud andmed on avalikud, siis võivad kihlveokontorid vähimagi kahtluse korral IP-aadressi ligipääsu piirata. See kehtib nii veebilehelt kraapimise kui ka serverist päringute tegemise korral.

Teiseks puudujäägiks võib lugeda kihlveokontorite veebilehtede ülesehituse erinevusi. Samuti on erinevad ka kihlveokontorite serverisse tehtavad päringud ning nende parameetrid. Iga erineva ülesehitusega veebilehelt kraapimiseks oleks vaja luua erinev programm, et koefitsiente kätte saada. Erinevad peavad olema ka kihlveokontorite serveritesse päringuid tegevad programmid. Kui mõni kihlveokontor oma veebilehe või päringute struktuuris midagi muudab, tuleb muuta ka vastavat programmi.

KOKKUVÕTE

Spordiennustus on tehing panustaja ja kihlveokontori vahel. Panustaja paneb raha mingi konkreetse sündmuse peale ning sündmuse toimumise korral saab tulu, mis on üheselt määratud panustatud summa ja kihlveokontori määratud antud sündmusele vastava koefitsiendi poolt. Kihlveokontorid võivad koefitsiente ise arvutada või osta neid ettevõtelt, kes nende arvutamise tegelevad.

Bakalaureusetöö üheks eesmärgiks oli kirjeldada koefitsientide müügiga tegeleva ettevõtte X olemasolevat manuaalset tööprotsessi. Hetkel on töötaja ülesandeks jälgida turu koefitsiente ning tagada turu keskmiste koefitsientide ja ettevõtte X koefitsientide arvutamiseks mõeldud programmis olevate koefitsientide vastavus. Selline tööprotsess toetub suurel määral inimese täpsusele ning keskendumisvõimele, mis omakorda võivad endaga kaasas tuua palju vigu.

Teiseks eesmärgiks oli kirjeldada potentsiaalseid võimalusi uue automatiseeritud süsteemi loomiseks, mis lihtsustaks töötajate tööd ning võimaldaks ettevõttel hallata korraga rohkemate spordivõistluste koefitsiente, kui see manuaalselt võimalik on. Töö põhiliseks tulemuseks oli kahe erineva võimaluse kirjeldused kihlveokontoritelt koefitsientide saamiseks. Samuti valmis ka mõlema võimaluse jaoks lühike prototüüp. Esimeseks võimaluseks oli koefitsientide kraapimine otse kihlveokontori veebilehelt, kasutades selleks programmi Selenium. Teiseks võimaluseks oli koefitsientide küsimine kihlveokontori serverist, saates sinna päringuid, mis küsivad mingi kindla mängu koefitsiente mingil kindlal ajahetkel. Esimene võimalus on küllaltki aeglane, sest veebilehed sisaldavad ka palju muid andmeid ja nende laadimine võib võtta kaua aega. Teist varianti kasutades saadakse serverist päringu vastuseks ainult vajaminevad andmed. Saadud koefitsiendid edastatakse programmile, mis sisendite põhjal arvutab turu keskmised koefitsiendid ning mis omakorda võib neid edastada ettevõtte X koefitsientide haldamiseks mõeldud programmile.

Töö tulemustest võib järeldada, et käesolevat ettevõtte X manuaalset koefitsientide haldamissüsteemi on võimalik automatiseerida, kasutades kahte kirjeldatud meetodit

koefitsientide saamiseks. Automatiseeritud tööprotsess toetub vähem inimvõimekusele ning on seeläbi efektiivsem ja kiirem. Samas võivad kihlveokontorid ise taolise tööprotsessi kasutamist piirata.

Käesoleva töö raames ei loodud täielikku automatiseeritud tööprotsessi. Töö võimaliku edasiarendusena on võimalik selline tööprotsess luua, kasutades töös kirjeldatud meetodeid.

SUMMARY

Sports-betting is a very popular means of entertainment as well as earning money. It is a transaction between a client or a bettor and a bookmaker. Bettor bets money on a certain event and if this event occurs they make profit. The size of the profit is fixed by the size of a bet and odds that are provided by the bookmaker for that specific event. Bookmakers can calculate their own odds or they can buy them from companies that calculate odds for them.

The first aim of the thesis is to describe a manual workflow of a company that provides this kind of a service to the bookmakers. Usually the worker keeps an eye on the odds on the market and inputs the necessary parameters to a specific program to produce new odds similar to the average of the market. This can be relatively slow and error-prone.

The second aim of this thesis is to describe two potential possibilities to automate current workflow. First, the odds can be scraped from the websites of the larger bookmakers, which contribute the most to the overall market. This can be implemented by using a website testing program Selenium. However, the faster method would be to request the odds from the bookmakers servers. The drawback is that the bookmakers often limit the access to their servers and websites if they feel threatened in any way.

As a result of this work, it can be concluded that it is possible to automate the current workflow by scraping odds from bookmakers websites and by requesting odds straight from bookmakers' servers. Automated workflow relies less on a worker's capabilities and is faster and more effective than the manual workflow.

KASUTATUD ALLIKAD

Anderson, P.M., Blackshaw, I.S., Siekmann, R.C.R., Soek, J. (2012). *Sports Betting: Law and Policy*. The Hague: Asser Press

Betting arbitrage. [WWW] <http://www.sportsbettingworm.com/arbitrage-betting/>
(17.05.2016)

Buchdal, J. (2003). *Fixed Odds Sports Betting*. London: High Stakes Publishing

Keogh, F., Rose, G. (2013). *Football betting – the global gambling industry worth billions*. [WWW] <http://www.bbc.com/sport/football/24354124> (10.05.2016)

Rovell, D. (2014). *Bettor cashes in on Luis Suarez bite*. [WWW] <http://www.espnfc.com/fifa-world-cup/story/1910428/norwegian-bettor-cashes-in-on-luis-suarezs-bite> (8.05.2016)

Korpela, J. (2003). *Methods GET and POST in HTML forms – what's the difference?*. [WWW] <https://www.cs.tut.fi/~jkorpela/forms/methods.html> (17.05.2016)

Selenium Master – How To Install Selenium Python Driver. [WWW] <http://seleniummaster.com/sitecontent/index.php/selenium-web-driver-menu/selenium-test-automation-with-python-menu/186-how-to-install-selenium-python-webdriver>
(15.05.2016)

SeleniumHQ Browser Automation. [WWW] http://www.seleniumhq.org/docs/01_introducing_selenium.jsp (15.05.2016)

LISA 1. SELENIUM

Käesolev lisa sisaldab Pythoni koodi kihlveokontori SBOBET veebilehelt koefitsientide kraapimiseks, kasutades programmi Selenium.

```
1. import time
2. import datetime
3. from selenium import webdriver
4. from selenium.webdriver.support.ui import WebDriverWait
5. from selenium.common.exceptions import StaleElementReferenceException
6. from selenium.common.exceptions import NoSuchElementException
7.
8. browser = webdriver.Firefox()
9. browser.implicitly_wait(10)
10.
11. url = "https://www.sbobet.com/euro/live-betting/football/estonia-premium-
    liga/1896017/flora-tallinn-vs-jk-nomme-kalju"
12. #käimasoleva mängu url sbobet.com leheküljelt, mille koefitsiente soovitakse uurida
13. browser.get(url)
14.
15.
16.
17. while True:
18.     values = []
19.     try:
20.         over_under = browser.find_element_by_id("od-ma-1-5")
21.         #otsib lehelt divi nimega od-ma-1-5, mis sisaldab mängu tulemuse andmeid
22.
23.         data = over_under.find_element_by_class_name("MarketBd")
24.         #otsib eelneva divi seest divi nimega MarketBd, mis sisaldab ainult koefitsiente
    ilma sündmuse pealkirjata
25.
26.         table = data.find_element_by_tag_name("tbody")
27.         #otsib tabeli sisu mis sisaldab koefitsiente
28.
29.         for row in table.find_elements_by_tag_name("tr"):
30.             for cell in row.find_elements_by_tag_name("td"):
31.                 values.append(cell.text.replace("\n", " "))
32.
33.
34.         time.sleep(5)
35.         browser.find_element_by_id('bu:od:re:0').click()
36.         #vajutab leheküljel olevat koefitsientide uuendamiseks mõeldud nuppu
37.         time.sleep(5)
38.     except StaleElementReferenceException:
39.         continue
40.     #kui tekib stale element error, proovib uuesti
41.     except NoSuchElementException:
42.         print("Game Over!")
43.         break
44.     #kui lehel enam otsitavaid andmeid ei ole, siis rohkem neid otsida ei püüa
45.
46.     if len(values) >= (3):
47.         #tabel peab koosnema vähemalt kolmest lahtrist
48.         print(datetime.datetime.now().strftime("%d-%m-%Y %H:%M:%S"))
49.         print(values)
50.     else:
51.         print("SUSPENDED!")
52. browser.close()
```

LISA 2. *GET* PÄRING

Käesolev lisa sisaldab Pythoni koodi kihlveokontori SBOBET serverist koefitsientide küsimiseks.

```
1. import urllib2
2.
3. url = "https://www.sbobet.com/en/data/event?ts=a9199619&tk=2873,2873,25,1,233,1989
4.     1,1900121,20160518,1,4,0,4&vs=0"
5.
6. #GET meetod
7. data = urllib2.urlopen(url).read()
8. print(data)
```