

10th MathSport International Conference 2023

Book of Abstracts

Corvinus University of Budapest, Budapest, Hungary

26–28 June 2023

Welcome Address

We are very glad to welcome you at Budapest for the 10th Mathematics in Sports International conference. It is our pleasure to host the conference this year after the successful previous ones in Manchester, Groningen, Leuven, Loughborough, Padova, Athens, and Reading. Since the first MathSport conference, a lot of research has been made on topics related to Mathematics and Sports and the interest has increased considerably. MathSport conference is a gathering of Academics and practitioners covering a wide range of sports and methodologies and we hope that this year the meeting will cover even more sports. We have attempted to arrange a wide range of talks during the three days of the conference.

A complete list of the abstracts of the papers to be presented at the conference can be found in this book.

We hope that you will enjoy the meeting.

László Csató

Chair of the Local Organising Committee



Scientific Committee

Dries Goossens (Ghent University, Belgium)

Phil Scarf (Cardiff University, United Kingdom)

Marco Ferrante (University of Padova, Italy)

Dimitris Karlis (Athens University of Economics and Business, Greece)

Ruud Koning (University of Groningen, Netherlands)

Stephanie Kovalchik (Victoria University, Australia)

Ioannis Ntzoufras (Athens University of Economics and Business, Greece)

Alun Owen (Coventry University, United Kingdom)

James Reade (University of Reading, United Kingdom)

Frits Spieksma (TU Eindhoven, Netherlands)

Ray Stefani (California State University, Long Beach, United States)

Local Organising Committee

Chair

László Csató

Institute for Computer Science and Control (SZTAKI); Corvinus University of Budapest

Members

Sándor Bozóki

Institute for Computer Science and Control (SZTAKI); Corvinus University of Budapest

Dóra Gréta Petrőczy

Institute of Economics, Centre for Economic and Regional Studies (KRTK KTI)
Corvinus University of Budapest

Zsombor Szádóczki

Institute for Computer Science and Control (SZTAKI); Corvinus University of Budapest

Contents

Welcome Address	i
Keynote talks	1
A flexible mixed model for the relationship between age and performance: Application to football, golf and snooker (<i>Ian McHale</i>)	1
The beautiful dataset (<i>Ignacio Palacios-Huerta</i>)	2
Optimizing soccer on-ball decisions: How to turn soccer players into agents? (<i>László Toka</i>)	3
Talks	4
Referee assignments in non-professional football leagues (<i>Tankut Atan</i>) . . .	4
Subject-specific workload management in rowing (<i>Arnold Baca</i>)	5
A vector embeddings representation of road cycling riders and races (<i>Ethan Baron</i>)	6
A novel method for player selection in T20 cricket (<i>Bhaskar Basu</i>)	7
Favoritism under multiple sources of social pressure (<i>Gábor Békés</i>)	8
Improving the fairness of constrained assignments: The role of draw order in the UEFA Champions League (<i>László Csató</i>)	9
Does hosting the Olympics mean more medals? Synthetic control analysis (<i>Gergely Csurilla</i>)	10
The impact of tournament design on match-fixing (<i>Karel Devriesere</i>)	11
Visual analysis of control area in badminton doubles using drone video dataset (<i>Ning Ding</i>)	12
Scheduling women's football in Belgium (<i>Dries Goossens</i>)	13
Tournament designs and schedules for a FIFA World Cup with 12 groups of 4 teams (<i>Mario Guajardo</i>)	14
2026 FIFA World Cup: What is the best format? (<i>Julien Guyon</i>)	15
Decision support for sports choices based on performance measurement using the generalized Thurstone method: the model and a case study (<i>László Gyarmati</i>)	16
Performance monitoring in anti-doping with Bayesian longitudinal models (<i>Laurentiu C. Hinoveanu</i>)	17
Classification of movements of soccer players using time-series data mea- sured by inertial measurement units (<i>Nobuyoshi Hirotsu</i>)	18
Does a superleague tournament kill domestic leagues? Evidence from the basketball's Euroleague (<i>Juan de Dios Tena Horrillo</i>)	19

Ranking handball teams using decision-making methods (<i>András Horváth</i>)	20
A machine learning based approach to analyse player performance in Twenty20 Cricket Internationals (<i>Gordon Hunter</i>)	21
Improvements in basketball modelling: some new models (<i>Dimitris Karlis</i>)	22
Market characterization of betting in basketball: Evidence from European competitions (<i>Christos Katris</i>)	23
The evaluation of the performance of golfers on LPGA tour based on LGCM approach (<i>Beom-Jin Kim</i>)	24
The fairness of berth allocation of the 48-team FIFA World Cup (<i>László Marcell Kiss</i>)	25
Scoring probability model based on service landing location and ranking points in men's professional tennis matches (<i>Eiji Konaka</i>)	26
Feedback, gender, and choking under pressure: Evidence from alpine skiing (<i>Alex Krumer</i>)	27
MIP formulations to create schedules for the new format of the UEFA Champions League (<i>Emiliano Lancini</i>)	28
Does draft currency promote competitive balance? An empirical investigation of the National Football League 2002-2021 (<i>Michael Lapré</i>)	29
A stochastic analysis of the 4 x 100 m relay (<i>Anders Lunander</i>)	30
Analyzing the profitability on the football transfer market using network science and machine learning (<i>Csaba Medgyes</i>)	31
Modelling women's football scores (<i>Rouven Michels</i>)	32
Sports wars-contest theory approach (<i>Yizhaq Minchuk</i>)	33
Predicting play-calls in the NFL using Markov-switching decision trees (<i>Marius Ötting</i>)	34
The probability of violating the Condorcet principle by Formula One points scoring systems (<i>Dóra Gréta Petrőczy</i>)	35
Pitcher selection scenarios in baseball: A multi-criteria approach (<i>Marina Polyashuk</i>)	36
Football group draw probabilities and corrections (<i>Gareth Roberts</i>)	37
Analysis of changes in curling after the introduction of five stone Free Guard Zone rule and No Tick Shot rule (<i>Alex Semenov</i>)	38
Spatial statistics of images from players point of view in sport environments (<i>Adam Shavit</i>)	39
Different rankings of European men's national football teams (<i>Vilmos Siteri</i>)	40
Application of genetic algorithm to solve a special team orienteering problem (<i>Benedek Soós</i>)	41
Round robin tournaments and integer programming formulations (<i>Frits Spieksma</i>)	42
Applications of point pattern analysis in baseball (<i>Dana Sylvan</i>)	43
Causal problems involving football strategy (<i>Tim Swartz</i>)	44
The design of costly rounds of paired comparisons: the GRAPH of regular graphs (<i>Zsombor Szádóczi</i>)	45
What do we not understand in (sport) rankings? (<i>Leszek Szczeciński</i>)	46
Defensive team analysis in the 2022 World Cup based on the event prediction (<i>Rikuhei Umemoto</i>)	47

What algorithm to select to create your sports schedule? (<i>David Van Bulck</i>)	48
Efficiency of betting markets in European football: Evidence from simulations and real data (<i>David Winkelmann</i>)	49
Profitable sports forecasting – Issues in evaluating predictive power and implications for model selection (<i>Fabian Wunderlich</i>)	50
Learning and planning to solve boulder climbing problems (<i>Barbaros Yet</i>) .	51
Prediction of shot type and hit location based on pose information using badminton match videos (<i>Tatsuya Yoshikawa</i>)	52
A dynamic, scalable algorithm to optimize the allocation of athletics scholarships (<i>Nizar Zaarour</i>)	53

Keynote talks

A flexible mixed model for the relationship between age and performance: Application to football, golf and snooker

Ian McHale

University of Liverpool

`ian.mchale@liverpool.ac.uk`

In this talk I will present a new mixed linear model for the relationship between age and performance. The model allows for random effects at the nodes of a barycentric interpolation, such that performance evolves with age in a non-prescriptive way. We use the model to investigate the effects of age on performance in football, golf and snooker - sports which have different physical and mental demands. When are you 'past it'? All will be revealed!

The beautiful dataset

Ignacio Palacios-Huerta

London School of Economics

`i.palacioshuerta@gmail.com`

Natural experiments, clean observability, precise measurement, high stakes, expert subjects, unimaginable detail, large datasets, exogenous rule changes, quasi-experimental variation, observable social effects, no Hawthorne effects. These and other desirable attributes for empirical work are found in sports settings. In sports, lab and field features often complement each other. Sports can offer the best of both worlds. A resistance to recognize these advantages reflects a misunderstanding of the virtues of sports data, and this resistance has discouraged the study of these settings and slowed down the production of knowledge. This survey reviews a literature that shows how sports settings have made possible the first successful test, or the best test to date, of various economic hypotheses, as well as the discovery of new phenomena. The perspective is not what economics can do for sports, but what sports can do for economics.

Optimizing soccer on-ball decisions: How to turn soccer players into agents?

László Toka

Budapest University of Technology and Economics

`toka.laszlo@vik.bme.hu`

Soccer is a complex sport that requires players to make real-time decisions while navigating a dynamic and uncertain environment. In recent years, there has been a growing interest in the use of machine learning techniques to help players improve their decision making on the field. However, the state-of-the-art mostly focuses on simulating soccer games or merely evaluating actions and players. Would it be possible to use soccer event and tracking datasets without any simulation (i.e., changing the historical actions) to derive the optimal on-ball actions, such as dribbling, passing direction, and shooting, for each time-step of the game in real-time? How about testing the outcome of the optimal strategy without the possibility of deploying it in a game? A robust optimization approach has the potential to revolutionize the way soccer is played and coached. Fortunately, recent breakthroughs in offline reinforcement learning are here to help us with these optimization tasks.

Talks

Referee assignments in non-professional football leagues

Tankut Atan

Bahçeşehir University

`sabritankut.atan@eng.bau.edu.tr`

Co-authors: Burak Çavdaroglu, Zühal Özcan Yavuz

The Referee Assignment Problem deals with the appointment of referees to a given tournament schedule. It is known to be an NP-hard problem. In non-professional football, the number of leagues and matches for which the referee assignments have to be decided can become quite large. For example, only in the city of Istanbul a committee has to make the referee assignment decisions for a total of 277 groups (tournaments) with 2271 teams. Solving a monolithic mixed-integer programming formulation exactly runs into memory and run-time problems. We report our results when the problem is partitioned into sub-problems which are solved iteratively.

Subject-specific workload management in rowing

Arnold Baca

University of Vienna

`arnold.baca@univie.ac.at`

Co-author: Philipp Kornfeind

Training impact analysis is a method of explaining the relationship between load and performance in training practice to support the optimization of the individual training process. The antagonistic “Performance Potential Double Model” (PerPot DoMo) considers both volume and intensity of the training load and is superior to related models using just one input quantity. Previous research in rowing has demonstrated its effectiveness in process-accompanying training control. Within the frame of a project intending to develop appropriate statistical approaches of machine learning for a novel modelling of the individual training process in elite rowing, we adapt and implement PerPot DoMo as a benchmark model. We investigate the influence of different combinations of loading parameters on the predicted performance using precisely measured performance parameters for validation. In addition, the new approach will also be tested in other endurance sports with regard to its general applicability.

A vector embeddings representation of road cycling riders and races

Ethan Baron

University of Toronto

`ethanbaron1105@gmail.com`

Co-authors: Bram Janssens, Matthias Bogaert

Vector embeddings have been successfully applied in various domains to obtain improved representations of non-numeric data. We present a novel application of vector embeddings in professional road cycling by constructing a bipartite graph with each node representing a specific rider or race, and each edge representing a rider's result in a given race. Then, we learn vector embeddings for each rider and race using node2vec, a representation learning technique for graphs. We use unsupervised learning techniques to show that the resultant embeddings capture interesting features of riders and races. Lastly, we explain how these embeddings could be used for downstream prediction tasks such as early talent identification.

A novel method for player selection in T20 cricket

Bhaskar Basu

XIM University

`bhaskar@xim.edu.in`

Co-author: Jimut Bahan Chakrabarty

Cricket as a sport has evolved over time and newer formats have emerged considering the changing tastes of the aficionados of the game. T20 cricket is the most recent and most popular international format of cricket. The fast-paced nature of the format demands teams to play with a different strategy and has scope for innovation compared to the longer formats of the sport. The commercial viability and brand value of the Indian Premiere League (IPL) has made the format popular among both players and spectators over the years. Under such circumstances, the traditional metrics to assess players' capabilities requires a fresh look. This paper aims to develop a novel method for player selection in T20 cricket by assessing players' performances in particular phases of the game using a proposed mathematical model. Adopting a dynamic approach of generating scores through the proposed model on a match-to-match basis, this article will help determine the best-equipped players on the field. The proposed model will be tested using most recent data from IPLT20 and international T20 cricket series played by India.

Favoritism under multiple sources of social pressure

Gábor Békés

Central European University
Institute of Economics, Centre for Economic and Regional Studies (KRTK KTI)

`bekesg@gmail.com`

Co-authors: Endre Borza, Márton Fleck

This paper shows that telling apart sources of social pressure is important. We revisit the empirical setting where football referees help the home team in vital situations. Contrary to previous hypotheses, this bias cannot be attributed to the presence of supporters in the stadium. Relying on exogenous variation in crowd presence due to Covid, we find that the bias persists without crowds. Instead, the alternative source of social pressure is the host team organization, with primarily the more influential (successful) teams benefiting from the referee bias. We also show that our findings are unlikely to be driven by corruption.

Improving the fairness of constrained assignments: The role of draw order in the UEFA Champions League

László Csató

Institute for Computer Science and Control (SZTAKI)
Corvinus University of Budapest

`laszlo.csato@sztaki.hu`

The organiser of the UEFA Champions League, one of the most prestigious football tournaments in the world, faces a non-trivial mechanism design problem each autumn: how to choose a perfect matching in a balanced bipartite graph randomly. For the sake of credibility and transparency, the Round of 16 draw consists of some discrete uniform choices from two urns whose compositions are dynamically updated with computer assistance. Even though the adopted mechanism is unevenly distributed over all valid assignments, it resembles the fairest possible lottery according to a recent result. We challenge this finding by analysing the effect of reversing the order of the urns. The optimal draw procedure is shown to be primarily dependent on the lexicographic order of degree sequences for the two sets of teams. An example is provided where exchanging the urns can reduce unfairness by one-third on average and almost halve the worst bias for all pairs of teams. Nonetheless, the current policy of starting the draw with the runners-up remains the best option if the draw order should be determined before the national associations of the clubs are known.

Does hosting the Olympics mean more medals? Synthetic control analysis

Gergely Csurilla

Institute of Economics, Centre for Economic and Regional Studies (KRTK KTI)

`csurilla.gergely@krtk.hu`

Co-author: Imre Fertő

In recent years empirical evidence has shown that host effect at the Olympics is not as obvious (Csurilla & Fertő, 2023), or at least that the magnitude of it is certainly reducing (Singleton et al., 2021). In addition to the host effect, pre- and post-host effects have also been shown previously, but the evidence for these is so far contradictory. With this paper, we want to provide evidence on whether the hosting of the event will indeed result in additional medals. Moreover, whether there is a demonstrable follow-up effect, a kind of medium- to long-term competitive advantage due to the additional resources invested in elite sport. We use the Synthetic Control Approach to demonstrate this. Kobierecki & Pierzgalski (2022) employed this method to investigate the impact of mega-sport events on tangible economic benefits. As in recent studies on Olympic success, we use sport-specific data to control for biases due to heterogeneities within sports.

The impact of tournament design on match-fixing

Karel Devriesere

Ghent University

`karel.devriesere@ugent.be`

Co-author: Dries Goossens

This presentation examines the relationship between tournament design and match-fixing. We evaluate various tournament design elements such as scheduling choices, ranking methods and tournament formats and their potential impact on match-fixing. Special emphasis is given to incentives that might arise to lose a game deliberately. Results from the literature indicate that certain tournament design elements can indeed increase the likelihood of match-fixing, particularly those that facilitate the occurrences of irrelevant matches and opportunities for collusion. While a major part of the literature focuses on match-fixing occurring in small tournaments with four teams, results from a wide variety of tournament sizes and formats will be discussed. Based on these literature findings, a set of rules is proposed for tournament organizers to minimize the risk of match-fixing occurring.

Visual analysis of control area in badminton doubles using drone video dataset

Ning Ding

Nagoya University

`ding.ning@g.sp.m.is.nagoya-u.ac.jp`

Co-authors: Yingjiu Bei, Keisuke Fujii, Kazuya Takeda

Visual performance analysis of players in a dynamic competition is vital for effective sports coaching. In racket sports, most previous studies focused on analyzing and assessing singles players and the discrete representations (e.g., stroke) that ignore meaningful spatial distributions. In our work, we built the first annotated drone dataset in badminton doubles and proposed a framework to estimate the control area probability map to evaluate teamwork performance in badminton doubles. In the experiment, we verified our approach by comparing various baselines. We also discovered the correlations between the score and control area and validated the effectiveness of our method through use cases. Our approach will provide coaches with valuable insights into doubles teamwork for assessing, recruiting, and scouting players.

Scheduling women's football in Belgium

Dries Goossens

Ghent University

`dries.goossens@ugent.be`

Co-author: Frits Spieksma

While the literature includes several papers on scheduling male football leagues with applications in various countries, women's leagues have hardly received attention. Due to lower interest by the general public (and hooligans), the scheduling process has probably not been perceived as very challenging or important. However, the success of the women's national team prompted a TV channel to live broadcast several matches in the highest division of the women's football league. This triggered a number of scheduling constraints which require a more sophisticated approach. For instance, while the TV channel could handle only one game per round, it wanted to broadcast a home game for each team at least once in the season. Moreover, as most teams lacked a suitable venue, live broadcasting was only possible when better infrastructure (typically from the men's teams) was available. We discuss the development of a tailor-made scheduling algorithm and its use in practice for the 2022-23 season.

Tournament designs and schedules for a FIFA World Cup with 12 groups of 4 teams

Mario Guajardo

NHH Norwegian School of Economics

`mario.guajardo@nhh.no`

Co-author: Alex Krumer

After the expansion of the FIFA World Cup from 32 to 48 teams starting in the 2026 edition, the initial proposal was to split the 48 national teams into 16 groups of three. Among other drawbacks, this proposal provides potential for collusion. Recently, after widespread criticism, FIFA officials decided to revisit that proposal, instead approving a tournament with 12 groups of four teams. However, the approved format does not eliminate the possibility of collusion. In this chapter, we propose tournament formats for a World Cup with 12 groups of four teams, considering several criteria, such as non-collusion, symmetry in rest days, and no dead rubbers. At the same time, our proposals attempt to adhere to the traditional format, with some nuances in either the group or elimination stage.

2026 FIFA World Cup: What is the best format?

Julien Guyon

Ecole des Ponts ParisTech

`julien.guyon@enpc.fr`

In 2018, we alerted FIFA and the general public that the 3-team group format that FIFA had planned for the 2026 World Cup had severe flaws: a high risk of collusion which could severely damage the image of soccer; schedule imbalance; and sporting and economic inefficiency, as $1/3$ of the teams are unnecessarily idle during the group stage [Guyon, “Why Groups of 3 Will Ruin the World Cup (So Enjoy This One)”, *The New York Times*, June 11, 2018; Guyon, “Risk of Collusion: Will Groups of 3 Ruin the FIFA World Cup?” *Journal of Sports Analytics*, 6(4):259–279, 2020]. In those articles we had also suggested better, alternate formats. Four years later, in 2022, FIFA finally seemed ready to move to a better format. In this talk, we compare the 16-groups-of-3, 12-groups-of-4, and 8-groups-of-6 (each made of two cross-competing groups of 3) formats, in terms of risk of collusion, financial revenue, and logistics (duration, number of stadiums, etc.).

Decision support for sports choices based on performance measurement using the generalized Thurstone method: the model and a case study

László Gyarmati

University of Pannonia

`xdkt dk96@gmail.com`

Co-authors: László Edvy, Csaba Mihálykó, Éva Orbán-Mihálykó

Paired/pairwise comparisons methods are frequently used in sports evaluations. One of them is the Thurstone method with more than two options. In this presentation, we extend the application of the method for decision support in sport selection. We present a model to compute weights for ranked assessment criteria based on coaches' opinions and, combined with the performance of prospective athletes who choose a sport in the assessment, to determine objectively and quantifiably which sport is most likely to be recommended to a given individual. This method allows us to measure the strength of each athlete objectively and compare them across age groups. Extending this method, it can be used not only for sport selection, but also for making various complex decisions when there are several possible outcomes to choose from. As the result of the research, we have generalized the set of criteria and applied it in a specific case study of athletics sport choice.

Performance monitoring in anti-doping with Bayesian longitudinal models

Laurentiu C. Hinoveanu

University of Kent

`L.Hinoveanu@kent.ac.uk`

Co-authors: Jim Griffin, James Hopker

In the fight against doping, there is an increasing need to develop methods which allow one to sensibly allocate testing resources. As the primary reason for doping is the improvement of athletic performance, it is reasonable to suggest that monitoring an individual's competition results on a longitudinal basis may reveal suspicious performance improvements. This work is an extension of a recently published performance model which aims to distinguish between normal or expected rates of progression and those caused by doping. We build a Bayesian spline model which also allows for skewed or heavy-tailed data. These assumptions lead to more robust estimators in the presence of poor performances. We find that athletes' trajectories follow a similar pattern, across performances in different sports measured by distance, time, or weight. We use our model to identify changes in the career performance trajectory of an athlete that may not be consistent with their age-matched cohort. These athletes can be flagged as individuals who may be at greater risk for doping and warrant follow-up investigation. We evaluate the performance of this approach on two data sets of athlete performances.

Classification of movements of soccer players using time-series data measured by inertial measurement units

Nobuyoshi Hirotsu

Juntendo University

`nhirotsu@juntendo.ac.jp`

Co-authors: Yuki Masui, Yu Shimasaki, Masafumi Yoshimura

This study aimed to classify the movements of soccer players during matches using raw data measured by inertial measurement units (IMUs). Twelve collegiate female soccer players were equipped with IMUs, and raw triaxial acceleration data from eight official matches were analyzed. The measurement data were separated for 3s, and a Fast Fourier Transform (FFT) was performed. After FFT, in order to classify the movements of players, we apply unsupervised clustering method of deep learning, aiming to group the FFT data into classes entirely without labels. Here, we try Invariant Information Clustering, which can be used to cluster any kind of unlabeled paired data by training a network to predict cluster identities. We also calculate the Euclidean distances between the data when the players were in the stationary state and other states using the k-means method. These methods could be used to classify the movements of not only collegiate female soccer players but also other soccer players.

Does a superleague tournament kill domestic leagues? Evidence from the basketball's Euroleague

Juan de Dios Tena Horrillo

University of Liverpool
University of Sassari

`jtena@liverpool.ac.uk`

Co-authors: Babatunde Buraimo, Jing Guan

While the impact of tournament designs on sports outcomes has received considerable research interest in the literature, less attention has been paid to studying how sports competitions interact. It is unclear how a strong European tournament affects the level of competition of national domestic leagues. We study the effect of participating in the basketball's Euroleague tournament on team performance in eight national leagues. We find that Euroleague participation harms national performance reducing the probability of victory by around 10%. However, this effect is only significant in big national leagues. A potential transmission channel of this effect is the reduction in the number of days between matches. Overall, the present paper suggests that a big continental competition can increase competitive balance at the national level.

Ranking handball teams using decision-making methods

András Horváth

Corvinus University of Budapest

horvathandras1218@gmail.com

In this study I use mathematical methods to create a power ranking of handball teams playing in the group stage of the 2022/23 season of the Men's EHF Champions League. While not all teams yet played against each other, I use variables such as goals scored, shot percentage, save percentage, number of turnovers, etc. to measure the effectiveness and performance of each team. I use the PROMETHEE decision-making method to evaluate the teams with respect to the criteria in the multi-criteria environment. Using the parameters from the PROMETHEE, I then also create a Bradley–Terry model to rank the teams, and to compare the results obtained from the different models.

A machine learning based approach to analyse player performance in Twenty20 Cricket Internationals

Gordon Hunter

Kingston University

`g.hunter@kingston.ac.uk`

Co-authors: Yash Jadwani, James Denholm-Price

Cricket is one of the most-followed sports around the World. T20 is a short version of the game, growing in popularity over the past 20 years due to high profile tournaments such as the Indian Premier League. There is much demand for analysis of player performance but traditional measures of this—batting and bowling averages, strike and economy rates—have limitations. We created a novel role-based performance metric using machine learning, allowing comparisons between players with similar roles in different teams. Using ESPNcricinfo data on T20 international matches, we calculated these new and the traditional performance metrics. Clustering was used to find “natural” classes of player types, a Random Forest classifier employed to identify the features most indicative of each cluster then PCA used to obtain the new performance indicator. Finally, we compared the results of our novel approach to the classical player performance metrics and player classifications provided by experts.

Improvements in basketball modelling: some new models

Dimitris Karlis

Athens University of Economics and Business

Karlis@aueb.gr

While there has been a lot of interest upon modelling the outcome (and in particular the score) of a basketball game, the approaches have been mainly restricted to bivariate Gaussian models. A reason for that is their simplicity for both fitting and deriving appropriate quantities for further evaluation/application of the models. In this paper, we will extend existing models by altering some of the basic assumptions. This is also based on data driven evidence from NBA results from different years, where the number of point scored by each team share interesting patterns. An intriguing such feature relates to the tail dependence, a high score for the one team is usually followed by a high score for the other teams perhaps explaining a fast pace of the game. We will present and apply a copula based approach to account for tail dependence and we will discuss the gain and the insights from this extended model.

Market characterization of betting in basketball: Evidence from European competitions

Christos Katris

University of Patras

`chriskatris@upatras.gr`

The scope of this paper is to investigate the nature of betting basketball markets in European competitions. Specifically, we consider the Euroleague 2017–2022 and the Eurocup 2017–2022 from European basketball. The main scope and the core of this paper is to analyze the fairness of the markets. The evidence is based on the Efficient Market Hypothesis (EMH) of Fama and is implemented through variance-ratio tests. Finally, some betting strategies (simple, from Modern Portfolio Theory and from the Risk Management domain) are developing and are considered as an example in order to test whether a player could make money from betting to these markets. The consequences of such a study are whether markets can be characterized according to fairness. Some betting strategies are testing with the use of real data from the European competitions and some conclusions about their effectiveness are extracted.

The evaluation of the performance of golfers on LPGA tour based on LGCM approach

Beom-Jin Kim

WCA Canada

`bkim25@wcacanada.com`

Co-author: Taeho Kim

Latent growth curve model (LGCM) is used to examine the performance of golfers on the LPGA tour. The panel data comes from the recent three seasons (2020, 2021, and 2022). The dependent variable is the official money list of players and the independent variables consist of continent, efficiency, technical factor (the average number of birdies, eagles, and holes in one), and mental factor (driving accuracy and putting average) of players. Each factor is determined by the weighted combination of parenthesized variables using Principal component analysis (PCA) and is measured as two latent variables of the intercept and the slope of the growth pattern for 3 years. Efficiency is evaluated by using Data envelopment analysis (DEA) bootstrapping with the output factors of scoring average and the rounds in the 60's and the input factors of Fried et al. (2004). Harvey Penick said that golf performance is dependent on a various factors and golf is psychology. These issues are investigated.

The fairness of berth allocation of the 48-team FIFA World Cup

László Marcell Kiss

Corvinus University of Budapest

kissmarci00@gmail.com

The new 48-team FIFA World Cup and the changes in berth allocation sparked debate among football fans, as the proportion of UEFA and CONMEBOL teams qualifying for World Cup decreased, while the rules highly favour AFC and CAF. This paper aims to add to the debate by investigating the fairness of distribution of berths between continents for the forthcoming World Cup. Fairness is meant as whether the number of berths allocated to continents reflect their past performance on World Cups. Performance shown at previous World Cups is a favourable basis for a ranking between continents, as World Cups are the most—and maybe the only—competitive form of intercontinental play. Pairwise Comparison Matrix based and Bradley-Terry methods are used on previous World Cup matches where different continents met to calculate their relative strength and the number of berths they deserve according to their strength. Several ways of robustness checks are performed on the results.

Scoring probability model based on service landing location and ranking points in men's professional tennis matches

Eiji Konaka

Meijo University

`konaka@meijo-u.ac.jp`

Co-author: Fumiya Shimizu

Service is regarded as the most crucial shot in tennis matches, affecting the outcome. The main objective of this research is to construct a mathematical model that expresses the scoring probability of service shots. The explanatory variables are the service landing location and the ranking points of the server and receiver (logarithms of ranking points for the two players). The model construction method is a stepwise regression of the generalized linear model (GLM). The model has been constructed using more than 110 thousand service shots from ATP Tour tournaments in 2019. The model consists of four service situations, 1st/2nd and ad/deuce sides. The model produced reveals the following facts: In 1st service, advantageous locations for servers are not symmetric between the ad and deuce sides. The higher-ranked player can win more points than lower-ranked players even if his shot lands in an easy location, i.e., in the middle of the service court.

Feedback, gender, and choking under pressure: Evidence from alpine skiing

Alex Krumer

Molde University College

`alex.krumer@himolde.no`

Co-authors: Christoph Bühren, Martin Gschwend

In alpine skiing competitions, one of the coaches of the participating countries sets the course. We analyze 40,150 men's and 36,968 women's performances from all competitions in alpine skiing's Slalom, Giant Slalom, and Super Giant disciplines that took place in the World Cups, World Championships, and Olympic Games between the 2001-02 and 2017-2018 seasons. We compare the performance of racers when competing on a course that was set by their compatriot to the performance of the same racers in the same season when the course was set by a coach from another country. Having a compatriot course setter only has an effect in the decisive run of the most technical discipline Slalom. We find that men fail significantly more often to complete their run when their compatriots set the course, whereas women fail significantly less in the same situation. The most likely drivers of our results relate to gender differences in response to feedback and choking under pressure in skill-based tasks.

MIP formulations to create schedules for the new format of the UEFA Champions League

Emiliano Lancini

Paris Dauphine University

`emiliano.lancini@lamsade.dauphine.fr`

Co-authors: Michele Barbato, Diego Delle Donne

The UEFA Champions League will adopt a new format starting with the 2024-25 edition. The number of teams in the competition will increase from 32 to 36, with 189 matches instead of 125, and the group stage will be replaced by a league phase following the so-called Swiss model. UEFA is expected to rank the teams in 4 pots of 9 teams each, based on a prespecified ranking. Each team will play against 2 teams from each pot (one home, one away), with the goal of creating an eight-match fixture list of roughly equal strength. We discuss the computational hardness of the problem of scheduling the UEFA Champions League's new format and propose integer programming formulations (IPF) to tackle it. To our knowledge, no previous IPF has been proposed for the Swiss model. Different constraints and objective functions are evaluated in the search of a balanced assignment of the matches considering several characteristics (e.g., travelled distance, balance on the average ranking of the rivals, etc.).

Does draft currency promote competitive balance? An empirical investigation of the National Football League 2002-2021

Michael Lapré

Vanderbilt University

`m.lapre@vanderbilt.edu`

Co-author: Elizabeth Palazzolo

In the National Football League (NFL) annual draft, teams take turns selecting entering players. The draft is a market mechanism designed to promote competitive balance as the NFL assigns draft positions to teams in reverse order relative to last season's performance. Teams frequently trade draft picks for other picks and/or players. We use several market valuations of draft picks to define original draft currency as the total value of draft picks available before any trades and final draft currency as the total value of picks used after all trades. For the 2002–2021 period, counter to the intent of the draft, we find that original draft currency does not affect the probability of reaching the playoffs but final draft currency does. Usage of outdated market valuations of draft picks by most teams can help explain how the best team has used draft-pick trades to remain among the strongest teams over two decades thereby perpetuating competitive imbalance.

A stochastic analysis of the 4 x 100 m relay

Anders Lunander

Örebro University

`anders.lunander@oru.se`

Co-authors: Niklas Karlsson

The baton exchanges are undoubtedly the most critical parts of the 4 x 100 meters relay race. Timing of the outgoing runner is critical. In this paper we analyze the race as a minimization problem under uncertainty. We formulate a stochastic model in which an outgoing runner at the baton exchange cannot perfectly assess the incoming runner's exact location relative a check mark position, and therefore potentially misjudges the right moment to start running. Also, the team members' daily shape is subject to uncertainty. To understand the effect of these two random variables—incoming runners' distance to checkmark and the daily shape of the running team—we conduct a simulation study to investigate the trade-off between the team's expected race time and their probability of being disqualified due to overrunning the takeover zone. The running profile of the relay team is taken from Ward-Smith and Radford (2002). Conditioning on a low disqualification probability, the difference in expected race time is shown to be substantial between teams with different variation in distance assessment and forecasting running performance, respectively.

Analyzing the profitability on the football transfer market using network science and machine learning

Csaba Medgyes

Budapest University of Technology and Economics

`csaba.medgyes@gmail.com`

Co-author: Roland Molontay

The transfer fees in football have skyrocketed in the last 20 years, making the transfer market a valuable source of income for clubs, to the extent that several teams base their business plan on maximizing profits from player transfers. However, little research has been conducted on football's transfer market from a network science perspective. In this work, the market is considered as a time-varying network where the nodes are the football clubs (or leagues) and the links represent transfer relationship between them with the transfer fees as weights. We collected more than 250,000 transfers from season 2010/11 until season 2021/22 considering the top-level leagues of 32 countries, and constructed a large, time-varying football transfer network. After visualizing the network and investigating its most important structural characteristics, we use machine learning methods to predict the profitability of a transfer by combining network attributes and player attributes.

Modelling women's football scores

Rouven Michels

Bielefeld University

`r.michels@uni-bielefeld.de`

Co-authors: Marius Ötting, Dimitris Karlis

For describing the number of goals in football, the model by Dixon and Coles (1997) has found tremendous impact. By extending the classical double Poisson model with two independent Poisson distributions in such a way that the probabilities for 0-0, 1-0, 0-1 and 1-1 can be changed, this model is widely considered as the standard model for football scores. In this talk, we show that this model is a special case of a multiplicative model known as the Sarmanov family. Moreover, we extend the classical Dixon and Coles model in various ways. Those extended models are then fitted to women's football data as previous models have been applied to men's football only. However, the scores in women's football are different to those of men's football and thus, an independent analysis is needed. We find that an extended Sarmanov model emerges as the most promising model for women's football scores.

Sports wars-contest theory approach

Yizhaq Minchuk

Shamoon College of Engineering

yizhami@sce.ac.il

We study cities' competitions for hosting sports teams or large-scale international sports events as a winner-pay contest with an entry fee. In the first stage, each contestant (city) decides whether to pay the entry fee (for example, infrastructure requirements); only by paying the first stage fee can the participant take part in the next stage, the contest stage. We show that when the contest organizer sets the optimal entry fee, the organizer's revenue does not depend on the number of contestants. Furthermore, in some cases, we have a form of the zero-sum game, when the sports team or international sports event organizer is the main beneficiary. In general, our study sheds additional light on this competition and shows that a city cannot benefit from hosting sports teams or large-scale international sports events.

Predicting play-calls in the NFL using Markov-switching decision trees

Marius Ötting

Bielefeld University

`marius.oetting@uni-bielefeld.de`

In the National Football League (NFL), several recent studies consider play-by-play data to predict a team's next play (either pass or run). Such studies often consider decision trees, which are useful in practice when working with coaches due to their simple interpretation. However, decision trees are designed for cross-sectional data, whereas most in-play data sets in sports have a time series structure and thus exhibit serial correlation. This talk also considers NFL play-by-play data and decision trees to predict play calls but we explicitly make use of the data's time series structure. We combine decision trees with hidden Markov models where an underlying (hidden) Markov chain selects the tree that generates the corresponding observation. Here, underlying states serve as proxies for a team's (latent) strategy, such as risky vs non-risky play. We propose an estimation approach based on the EM algorithm and use 8 seasons of NFL data to demonstrate the usefulness of our method.

The probability of violating the Condorcet principle by Formula One points scoring systems

Dóra Gréta Petróczy

Institute of Economics, Centre for Economic and Regional Studies (KRTK KTI)
Corvinus University of Budapest

`petroczy.dora@krtk.hu`

Co-authors: Attila Tasnádi, László Csató

Points scoring systems, widely used in multi-race sports championships, do not fulfil the requirement of Condorcet consistency, i.e., in some cases the Condorcet winner does not win the championship even though outperforming any other contestant in more than half of the races. More generally, it may be the case that one contestant is clearly better than another based on pairwise comparisons, but still scores fewer points. Our research investigates the frequency of this phenomenon based on the results of the Formula One racing series between 1973 and 2022.

Pitcher selection scenarios in baseball: A multi-criteria approach

Marina Polyashuk

Northeastern Illinois University

`m-polyashuk@neiu.edu`

This study focuses on various scenarios that are typical for a Major League Baseball (MLB) manager's routine decisions in pitcher selection. We consider three major reasons for such selections: rotational selection, relief selection, and special purpose selection. In various types of scenarios, we combine practical reasoning and available statistical data with the application of multi-criteria decision analysis to create mathematical models. The suggested models and methodologies are aimed at informing, supporting, and simplifying baseball manager's decisions in pitcher selections. Throughout the paper, the proposed methodologies are applied to the scenarios from the 2022 season of the Chicago White Sox team, and actual managerial decisions are compared to the solutions proposed by our models.

Football group draw probabilities and corrections

Gareth Roberts

University of Warwick

`gareth.o.roberts@warwick.ac.uk`

Co-author: Jeff Rosenthal

This talk will consider the challenge of designing football group draw mechanisms which have the uniform distribution over all valid draw assignments, but are also entertaining, practical, and transparent. Although this problem is trivial in completely symmetric problems, it becomes challenging when there are draw constraints which are not exchangeable across each of the competing teams, so that symmetry breaks down. The talk will explain how to simulate the FIFA Sequential Draw method, to compute the non-uniformity of its draws by comparison to a uniform Rejection Sampler. It will then propose two practical methods of achieving the uniform distribution while still using balls and bowls in a way which is suitable for a televised draw. The solutions can also be carried out interactively. The general methodology provided can readily be transported to different competition draws and is not restricted to football events.

Analysis of changes in curling after the introduction of five stone Free Guard Zone rule and No Tick Shot rule

Alex Semenov

HSE University

`semenoffalex@gmail.com`

Co-authors: Gleb Vasiliev

The project is dedicated to the study of curling. Over the past 5 years, the International Curling Federation has introduced two new rules. The five stone Free Guard Zone rule removes the option to take out guards before the 6th shot. The No Tick Shot rule prohibits both teams from removing guards from the center line in the early game, allowing the proper center setup. Of the expected improvements, a more interesting attacking game, an increase in the likelihood of a comeback, and a decrease in the number of blank ends were named. The authors want to understand how the expected effects of the introduction of the rules coincided with the actual change in the game.

Spatial statistics of images from players point of view in sport environments

Adam Shavit

City University of New York

`as1127@hunter.cuny.edu`

Co-authors: Dana Sylvan

Perception of the space around oneself is crucial for playing sports. First, to orient oneself relative to gravity, and then to interact with the environment. This work examines the visual spatial statistics available in images of sport related environments, for example large soccer arenas, and basketball courts, from the point of view of players in order to compare and contrast the visual information available to players with that available in other visual environments. The emphasis is on estimating spatial coordinates from this visual information. Specifically, we focus on spatial statistics for estimating surface slant and tilt, and the orientation spectrum at different spatial frequencies. We vary the type of sport venue, as well as the orientation of the camera, in order to generate and later analyze the spatial statistics of these scenes. We discuss the visual information available to players in sports related environments.

Different rankings of European men's national football teams

Vilmos Siteri

Corvinus University of Budapest

`vilmos.siteri@stud.uni-corvinus.hu`

This paper ranks the European men's national football teams and compare the results with the official ranking of FIFA and also with the outcome of the latest UEFA Nations League. In order to create a ranking, first the Pairwise Comparison Matrix (PCM) of the national teams is built, and the LLSM weight calculation technique is applied to gain the ranking. Furthermore, different robustness tests have been carried out, e.g., changing the LLSM to the EV method or creating different partial rankings. The relevancy of the study is supported by the recently finished World Championship and some placements of the official ranking, which are questioned by many (for instance, Belgium's first place).

Application of genetic algorithm to solve a special team orienteering problem

Benedek Soós

University of Pannonia

`soos2002.benedek@gmail.com`

Co-author: Anna Ibolya Pózna

In this paper, a special kind of team orienteering problem (TOP) is presented, that is a mixture of regular and score type orienteering courses. In this problem, the team has to complete a fixed-order course while visiting a set of elective control points too. The task of the team is to distribute the elective control points among the members in order to minimize the time required to complete the course. In this paper, an effective genetic algorithm based solution is proposed to solve this problem. Different variants of the genetic operators are implemented and their effect on the algorithm's performance are analysed. The solutions provided by the proposed algorithm are evaluated based on the results of real team orienteering competitions

Round robin tournaments and integer programming formulations

Frits Spieksma

Eindhoven University of Technology

`f.c.r.spieksma@tue.nl`

Co-authors: Jasper van Doornmalen, Christopher Hojny, Roel Lambers

Round robin tournaments are omnipresent in sport competitions and beyond. We propose two new integer programming formulations for scheduling a round robin tournament, one of which we call the matching formulation. We analytically compare their linear relaxations with the linear relaxation of a well-known traditional formulation. We find that the matching formulation is stronger than the other formulations, while its LP relaxation is solvable in polynomial time. In addition, we provide an exponentially sized class of valid inequalities for the matching formulation. Complementing our theoretical assessment of the strength of the different formulations, we also experimentally show that the matching formulation is superior on a broad set of instances. Finally, we describe a branch-and-price algorithm for finding round robin tournaments that is based on the matching formulation.

Applications of point pattern analysis in baseball

Dana Sylvan

City University of New York

`dsylvan@hunter.cuny.edu`

Co-author: Hisanobu Kaji

Statistics is an integral part of baseball, all the more in the last years with the huge developments in technology and data science. A wealth of baseball data is freely available and contains spatial information connected to location coordinates for individual batted balls captured by high-speed cameras. In this project we focus on modeling and visualization of baseball data via point pattern analysis. We use kernel smoothing to produce spatial heat maps of batted balls for top Major League Baseball players. In particular, we compare and contrast patterns produced by right-handed, left-handed, and switch hitters, respectively. After rejecting the hypothesis of complete spatial randomness we employ kernel smoothing to model and visualize the space-varying intensity function of the resulting inhomogeneous Poisson process. The procedure includes a routine for the selection of optimal smoothing parameters and the corresponding heat maps accurately identify the unique patterns in which players hit the baseball on the baseball field.

Causal problems involving football strategy

Tim Swartz

Simon Fraser University

`tim@stat.sfu.ca`

Co-authors: Nirodha Epasingheghe Dona, Lucas Wu, Joan Hu, Aaron Danielson

This paper investigates three problems involving strategy in football: (1) whether to cross the ball; (2) whether teams should play with pace; and (3) what are optimal throw-in locations. The problems are addressed using tracking data which facilitates the identification and quantification of confounding variables that are related to the outcome variables and the treatment variables. Although the data structure differ in the three problems, the line of research provides a template for addressing causal inference in sport. From a practical viewpoint, we obtain some surprising results.

The design of costly rounds of paired comparisons: the GRAPH of regular graphs

Zsombor Szádoczki

Institute for Computer Science and Control (SZTAKI)
Corvinus University of Budapest

`szadoczki.zsombor@sztaki.hu`

Co-author: Sándor Bozóki

Determining the pairs to be compared with each other in the consecutive rounds is a key question in preference modeling, decision making and sports. We focus on the graph representation of incomplete pairwise comparison matrices. Based on several numerical simulations, we propose a design of the rounds in the case of different number of teams in a way that the gained logarithmic least squares ranking vector is always the closest (according to the Euclidean distance) to the one that is gained from the total round-robin tournament. The findings are visualized with the help of a GRAPH of regular graphs. Although the results can be used in a much wider field, one of the main applications can be the case, when in a sports tournament, the teams' abilities are not known a priori, and organizing rounds has a significant cost.

What do we not understand in (sport) rankings?

Leszek Szczecinski

Institut National de la Recherche Scientifique

`leszek.szczecinski@inrs.ca`

Ranking of the teams in sports is one of the fundamental issues where the most popular rule consists in summing the points associated with the wins and the draws of each team. Teams are then ranked by sorting the score obtained. There are two main issues with this rule. First, the ranking results do not have a clear interpretation, and, second, they depend on subjectively defined value of the wins. For example, in association football games, win is worth three points, and, in ice hockey ranking, two. Our goal is to show that this value can be objectively determined from the data. Using mathematical model relating the order between teams to the outcome of the games, we show that summing the points of wins and draws is a ranking rule consistent with the model. We then explain how to estimate the parameters of the model, from which the value of win is obtained. As a practical example, we show that data indicates that the value of the win in football (three points) is too small.

Defensive team analysis in the 2022 World Cup based on the event prediction

Rikuhei Umemoto

Nagoya University

`umemoto.rikuhei@g.sp.m.is.nagoya-u.ac.jp`

Co-authors: Keisuke Fujii, Kazushi Tsutsui

Analyzing defenses in team sports is challenging because of the limited event data. Researchers have previously proposed methods to evaluate football team defense by predicting the events of ball gain and being attacked. However, they did not consider the importance of the events, and did not fully investigate the influence of the diversity (e.g., nationality). Here, we propose a generalized valuation method of defensive teams by score-scaling the predicted probabilities of the events. Using the location data of all players in football games of the 2022 World Cup, we investigated the effect of the number of players on the prediction and validated our approach by analyzing the games. Results show that for the predictions of being attacked, scoring, and conceding, all players' information was not necessary, while that of ball gain required information on three to four offensive and defensive players. We also explained the excellence in defensive teams in the 2022 World Cup.

What algorithm to select to create your sports schedule?

David Van Bulck

Ghent University

`david.vanbulck@ugent.be`

Co-authors: Dries Goossens, Jan-Patrick Clarner, Angelos Dimitzas, George H. G. Fonseca, Carlos Lamas-Fernandez, Antony E. Phillips, Roberto Maria Rosati

In a round-robin tournament, each team meets every other team exactly once. Scheduling real-life sports competitions turns out to be surprisingly difficult, hence various algorithms have been proposed over the years. Well-known techniques include integer-programming, meta-heuristics, or even satisfiability solvers. It turns out, however, that there is no single algorithm that is most suited to schedule all kinds of sports competition. Hence, the question arises what algorithm from the literature an organizer should buy or implement to schedule a number of future seasons of the sports competition at hand. This talk investigates the use of machine learning techniques to predict algorithm performance, when given the properties of the sports competition to be scheduled. In particular, we identify the properties that are most important to make that prediction, provide insights into the performance of existing scheduling algorithms, and provide suggestions to further improve them.

Efficiency of betting markets in European football: Evidence from simulations and real data

David Winkelmann

Bielefeld University

david.winkelmann@uni-bielefeld.de

Co-authors: Marius Ötting, Christian Deutscher

Previous literature leaves the impression that betting market inefficiencies are widespread. Since most studies cover relatively short periods, the power of hypothesis tests is potentially limited. It remains to be investigated whether market inefficiencies are systematic or whether their appearance is of random nature. We provide simulation-based analyses on betting markets, deriving the probability of detecting periods of inefficiencies in theoretical settings. This allows us to discuss whether the number of inefficient periods obtained in an empirical analysis of the top European football leagues exceeds what would be expected by chance only. Our results show that most inefficiencies obtained in real-world betting markets are short-lived and do not occur persistently over time or systematically across leagues. At the same time, the simulation-based analyses reveal that markets temporarily record inefficiencies even when full market efficiency is assumed.

Profitable sports forecasting – Issues in evaluating predictive power and implications for model selection

Fabian Wunderlich

German Sport University Cologne

`f.wunderlich@dshs-koeln.de`

Co-authors: Marc Garnica Caparrós, Daniel Memmert

In light of the multitude of perspectives on sports forecasting, it is not clear from the outset what makes a sports forecasting model a “good” model. The present contribution focuses on two of the most important goals in forecasting, namely accuracy and profitability. Based on theoretical considerations, artificial data generated with a tailored simulation-framework and real-world datasets, we uncover differences between model accuracy and model profitability, discuss resulting issues for evaluating predictive power and propose possible consequences for forecasting including model selection. While accuracy and profitability are related concepts, results illustrate that using profitability measures instead of accuracy measures to evaluate the predictive power of models has some serious pitfalls. These include sample size reduction, the existence of profitable but highly inaccurate models and possibly non-optimal choice of model parameters.

Learning and planning to solve boulder climbing problems

Barbaros Yet

Middle East Technical University

`byet@metu.edu.tr`

Co-authors: Kerem Demirtaş, Mert Türedioğlu

Boulder climbing is a well-balanced sport between physical, technical, and mental skill requirements as it tests athletes' decision-making and problem-solving abilities alongside their physical fitness. In climbing competitions, each climbing route is a novel decision problem that an athlete needs to solve in a limited amount of time. The athlete needs to plan the sequence of movements for a route they see for the first time before starting to climb. This study presents a goal-based climbing agent that learns from the previous solutions it observes to plan the sequence of actions for novel boulder climbing problems it encounters. The agent learns the cost of their movements from a dataset of the difficulty estimations and solutions provided by expert climbers and uses this information to solve the novel climbing problems it encounters. It also uses expert knowledge to constrain the learning and state space.

Prediction of shot type and hit location based on pose information using badminton match videos

Tatsuya Yoshikawa

Nagoya University

fujii@i.nagoya-u.ac.jp

Co-authors: Kazuya Takeda, Keisuke Fujii

With the development of video processing technology, it is expected to be used for various motion analysis in racket sports. The previous badminton model predicted stroke types and hitting positions based on match situations and player position information. However, prediction without motion information is difficult to apply to actual play. In this study, we propose a multi-task deep learning model that predicts the stroke type and hitting location based on the estimated pose information from badminton singles match videos. Experimental results show that the proposed method predicted the stroke type and hitting position more accurately than the existing method without pose. Furthermore, by introducing a mechanism called a gradient reversal layer to decrease the bias between the front and the back of the image, the proposed method predicted more accurately the stroke type. Our approach is expected to be used in games and practices by understanding easily predictable movements.

A dynamic, scalable algorithm to optimize the allocation of athletics scholarships

Nizar Zaarour

Northeastern University

`n.zaamour@northeastern.edu`

Co-authors: Basma Bargal, Emanuel Melachrinoudis

When a high school student athlete makes the decision to swim collegiately, one of the leading factors in choosing a university is the amount of the scholarship award. These awards are determined by the equivalency category in the National Collegiate Athletic Association (NCAA) Divisions I and II, and the National Association of Intercollegiate Athletics (NAIA). Rules govern financial aid and the number of scholarships allowed by a member institution. The exact financial distribution among athletes, however, is mostly the coaches' responsibility and there is scarcity in the literature on how these scholarships are distributed among athletes in a college/university, which it is important for athletics departments within these colleges and universities to have a model to help distribute the scholarship amounts in an optimal manner that promotes transparency and ethical leadership. This paper provides a model to aid coaches in the distribution of financial award/scholarships in women's swimming in a major urban university; however, the model can be applied and adjusted for financial award distribution for any collegiate sport following any institute specific policies and preferences.