IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

The role of the network of matches on predicting success in table tennis

This is the author's manuscript

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1669672 since 2018-11-07T13:25:06Z

Published version:

DOI:10.1080/02640414.2018.1482813

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

20 April 2024

The role of the network of matches on predicting success in table tennis

Mirko Lai[a,b] and Rosa Meo[a] and Rossano Schifanella[a] and Emilio Sulis[a]

[a]Computer Science Department, University of Torino, Italy; [b]Departamento de Sistemas Informàticos y Computación, Universitat Politécnica de Valéncia, Spain

ARTICLE HISTORY
Compiled November 6, 2018

ABSTRACT

The influence of training, posture, nutrition or psychological attitudes on an athlete's career is well described in literature. An additional factor of success that is widely recognized as crucial is the network of matches that an athlete plays during a season. The hypothesis is that the quality of a player's opponents affects her long-term ranking and performance. Even though the relevance of these factors is widely recognized as important, a quantitative characterization is missing. In this paper, we try to fill this gap combining network analysis and machine learning to estimate the contribution of the network of matches in predicting an athlete's success. We consider all the official games played by the Italian table tennis players between 2011 and 2016. We observe that the matches network shows scale-free behavior, typical of several real-world systems, and that different structural properties are positively correlated with the athletes' performance (Spearman $\rho = 0.88$, p-value < 0.01). Using these findings, we implement three different tasks, such as talent identification, performance and ranking prediction. Results shows consistently that machine learning approaches are able to predict players' success and that the topological features play an effective role in increasing their predictive power.

KEYWORDS
Social Network Analysis - Machine Learning - Performance Analysis - Talent Identification - Ranking Prediction

Introduction

Usually sport performance studies focus on aspects such as training activities, nutrition, behavior, and cognitive strategies. See for instance the studies by Lames and McGarry (2007); Nevill, Atkinson, and Hughes (2008). Experts widely acknowledge that self-improvement directly depends on the competitions and opponents a player is exposed to. In fact, by dealing with more qualified opponents and variegated styles of play, an athlete tends to improve his own experience and competitive strengths. However, it is difficult to empirically prove and quantify these intuitions. This study tries to fill this gap, analyzing how the topological properties of the graph of matches, i.e., who plays with whom, affects performance.

We focus on table tennis, one of the most practiced sports worldwide, and we consider a large dataset containing all table tennis games played in Italy between 2011 and

CONTACT E. Sulis. Email: sulis@di.unito.it

2016, as described in Sulis, Meo, Schifanella, and Arisi (2017). Matches are modeled as a graph where nodes are players and an edge between two nodes exists if they played against each other. The main contribution of this study is twofold: (1) we show how a machine learning framework can successively predict future talents and, in general, the success of players; (2) we quantitatively assess the importance of the topological properties of the network of matches in predicting success.

To introduce the argument, we briefly review some related works. Several studies on performance successes in sports largely investigated psychological aspects. See the work by Lejeune, Decker, and Sanchez (1994). Walling, Duda, and Chi (1993) focused on the role of motivation, while Ericsson, Krampe, and Tesch-Römer (1993) focused on deliberate practice. According to Woodman and Hardy (2003), anxiety has a relevant role as well as coping (see Nicholls and Polman (2007)). Other works, such as Cronin and Sleivert (2005), considered the role of training while Kusubori, Yoshida, and Sekiya (2012); Raab, Masters, and Maxwell (2005) considered the role of specific skills in table tennis. Many other insights concern mostly athletes' behavior: Atkinson and Reilly (1996) considered the role of monitoring the sleep cycle or Kondric, Sekulic, and Mandic (2010) considered the doping attitudes. Similarly, talent identification researches mainly investigate personal characteristics. For instance, Woods, Raynor, Bruce, McDonald, and Robertson (2016) performed a multi-dimensional assessment, combining the outcomes of tests on physical, technical and perceptual-cognitive performance. More generally, few works adopted a network perspective to study the athletes playing performance. The lack of a complete dataset as well as the lack of powerful computational capabilities in the management of large data can be a possible explanation.

Social Network Analysis (SNA) in sports has been already used by Girvan and Newman (2002) in the United States college football league with the purpose to automatically detect communities (i.e. groups of teams in the league conferences). Park and Newman (2005) proposed a sport ranking system based on network properties, while Fast and Jensen (2006) modeled the National Football League as a network to analyze the influence of notable coaches as well as to predict which teams will make the playoffs. The whole network of matches was described by Onody and de Castro (2004), who observed the increment of players' connectivity and the decrease of clustering coefficient, as well as by Radicchi (2011) that performs a complex network analysis of tennis matches played by professional players. Similarly to these studies, our work considers a large dataset of matches, as well as SNA metrics. Finally, the prediction of sports outcomes is relevant to address investment and reduce costs as observed by Pion, Hohmann, Liu, Lenoir, and Segers (2017). While efforts can be made to model the outcome of matches as done by McGarry and Franks (1994) and Ben-Naim, Redner, and Vazquez (2007), our interest here is mainly on the athletes' performance improvement and position in ranking. This is similar to the goal of Kovalchik (2016).

A table tennis match ends when a player wins the majority of sets. A set game is played up to 11 points, unless both players score 10 points. In this case, a game will be won by the first player to gain a two point lead. While some particular matches of final championships are played at best-of-five sets, usually matches are a best-of-three sets format. The Italian federation collects the results in a database which includes other information about athletes and tournaments. To compute the final ranking, the system is not trivial: each player has a score, which is assigned at the beginning of a season on the basis of the final position in the last season. The range of the scores varies proportionally from 15,000 points (the first athletes in the ranking) to 500 (the last ones). After each match, competitors change their score depending on the results: for

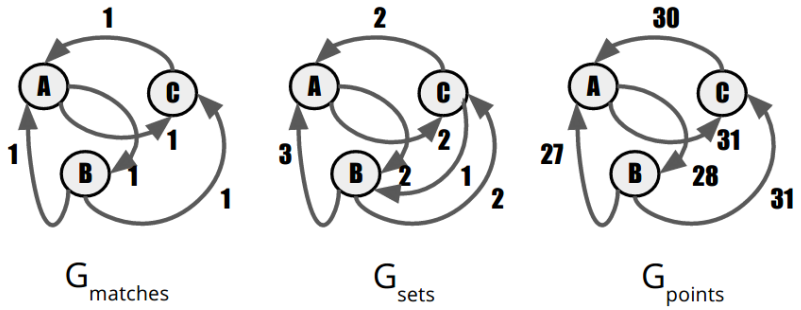

Figure 1. Example of the three different graphs corresponding to the activities in Table 1 when applying the $G^s_{matches}$, G^s_{sets}, and G^s_{points} weighting schemes.

Table 1. Simplified example of the games history between 3 players A, B, and C in a season.

Match	Winner	Sets	Points		
A vs B	A	2-0	11-8	11-9	
A vs C	A	2-0	11-5	11-9	
C vs B	C	2-1	12-10	5-11	11-4
A vs B	B	1-2	11-13	11-5	8-11
A vs C	C	0-2	5-11	8-11	

instance, winning against a player whose score is higher implies gaining more points. The final ranking of a season is computed on the basis of the points won and lost by athletes in the season.

Methods

Dataset

The dataset contains the complete set of 723,057 table tennis matches played by 21,458 players in Italy in five seasons S, from 2011/12 to 2015/16. For each season $s \in S$, we created three directed weighted graphs $G^s_{matches}$, G^s_{sets}, and G^s_{points} where the direction of a generic edge $A \xrightarrow{\text{w}} B$ goes from the defeated (A) to the winner player (B) and the weight w is computed with the following strategies:

$G^s_{matches}$
: number of matches won by B against A in season s.

G^s_{sets}
: number of sets won by B against A in season s.

G^s_{points}
: difference between the points gained by A and twice the points gained by B computed over all matches played between A and B during the season s.

Figure 1 shows an example of the different weighting schemes in a real scenario. It is worth noting that the choice of the weighting scheme has an impact in modeling the tie strength of a dyad: in fact, given the set of matches in Table 1, A and B played twice and each of them won one game, that corresponds to a mutual edge with

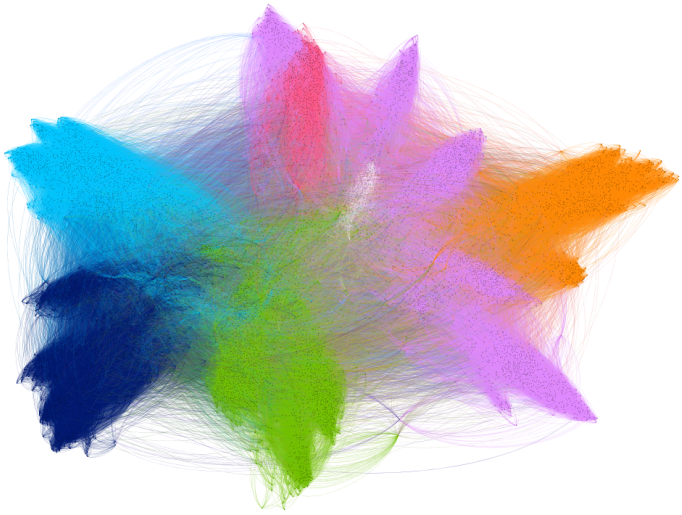

Figure 2. Graph of table tennis matches played in Italy from 2011/12 to 2015/16. Colors of vertices describe the communities detected with the *modularity* algorithm as implemented by the visualization tool Gephi[2]

the same weight $w_{A\to B} = w_{B\to A} = 1$ according to the $G^s_{matches}$ scheme. However, considering the G^s_{sets} scheme, the weights get unbalanced with A winning more sets, i.e., $w_{A\to B} < w_{B\to A}$ ($2 < 3$ sets). The opposite situation occurs if we consider the G^s_{points} strategy; in fact, $w_{A\to B} > w_{B\to A}$ ($28 > 27$ points).

For each weighting scheme, the graphs $\overline{G_*} = \bigcup_{s\in S} G^s_*$ model the aggregation across seasons of the different matches networks with $* \in \{matches, sets, points\}$.

Network metrics

In order to capture the role of the networks topology, in this work we computed a set of widely-used metrics, both at the graph-level and at the level of the nodes. Table 2 summarizes the meaning of the most important network metrics. As regards the graph-level measures we computed graph density, assortativity, average path length, and clustering coefficient. As regards the measures at the node-level we computed node degree and some node centrality measures such as betweenness, closeness, eigenvector, and PageRank. Some of these graph measures and (average) node measures are reported in Table 3 which summarizes the main structural properties for each of the seasonal networks $G^s_{matches}$, as well as for the aggregated network version $\overline{G_{matches}}$.

Independently of the weighting scheme, $\overline{G*}$ falls into the category of small-world networks in which most nodes can be reached from every other node via a small number of hops, i.e., low average path length (3.4) and it shows high average clustering coefficient (0.35) in comparison to a random graph with the same number of nodes and edges (0.017). Edge density is consistently low (0.0012). $\overline{G*}$ degree distribution follows a power-law tendency with few players highly connected (athletes which played hundreds of matches) and a long tail of less active amateur players. Finally, a consistent positive assortativity (0.3) indicates how athletes tend to play against players with similar activity pattern: this is coherent with the inclusion into categories of specific tournaments and championships. Values in parenthesis refer to the case of $\overline{G_{matches}}$.

Graph metrics	Definition	Bibliography
Degree, Indegree, Outdegree	degree is the number of edges incident to a node. In a directed graph, the number of incoming edges in a node is indegree; the number of outcoming edges is outdegree	Albert and Barabási (2002)
Average neighbors degree	average degree of the neighbors of a node	M. E. Newman (2008)
Density	ratio of the number of edges to the number of possible edges	Wasserman and Faust (1995)
Assortativity	measures the preference of a node to connect to similar nodes	M. E. J. Newman (2002)
Clustering coefficient	fraction of the possible triangles that exist through a node	M. E. Newman (2008)
Average path length	average number of steps along the shortest paths for all possible pairs of network nodes	Albert and Barabási (2002)
Betweenness centrality	number of shortest paths that pass through a node; it is similar to load centrality	M. E. J. Newman (2001)
Eigenvector centrality	measures the influence of a node in a network	M. E. Newman (2008)
Closeness centrality	inverse of the sum of the shortest distances between each node and every other node in the network	Bavelas (1950)
Degree centrality	centrality indicator that identifies the most important nodes on the basis of the node degree	M. Newman (2010)
PageRank centrality	node importance in a directed network, measured on the number of links in which the links that weigh more come from important nodes, with few links	Brin and Page (1998)

Table 2. Definition of the main network metrics used in this work

Season	**V**	**E**	AvDg	ClCo	Dens	AvPL
$G^{2011/12}_{matches}$	9,569	109,857	22.96	0.345	0.0025	3.84
$G^{2012/13}_{matches}$	9,851	113,285	23.00	0.344	0.0024	3.85
$G^{2013/14}_{matches}$	9,909	116,529	23.52	0.342	0.0024	3.89
$G^{2014/15}_{matches}$	9,883	118,794	24.04	0.337	0.0023	3.92
$G^{2015/16}_{matches}$	9,906	120,791	24.39	0.342	0.0024	3.94
$\overline{G_{matches}}$	21,458	558,482	61.25	0.354	0.0012	3.41

Table 3. Number of players (vertices, V), links (edges, E), average degree (AvDg), density (Dens), clustering coefficient (ClCo) and average path length (AvPL), for five seasonal networks $G^{s}_{matches}$ and their aggregated version $\overline{G_{matches}}$. A summary of the network metrics definitions can be found in Table 2.

Type	Name	Description
I	Age, Gen	Age and gender
	CityB,ProvB,RegB	City, Province and Region of birth
Q	N-Mat	Number of matches played
	W-Mat,W-Set,W-Poi	Number of won matches, sets, and points
	L-Mat,L-Set,L-Poi	Number of lose matches, sets, and points
	D-Mat,D-Set,D-Poi	Difference between the number of won and lose matches, sets, and points
T	Deg,De-In,De-Out	Degree, in-degree, out-degree
	C-Bet,C-Clo,C-Eig	Betweenness, closeness, eigenvector centrality
	PageRank,C-Loa,C-Deg	Pagerank, load, degree centrality
	AvND	Average neighbor degree

Table 4. Three sets of features for each player in the season including demographic (I), quality (Q), and topological (T) features.

Features

We model players' performance prediction as a machine learning problem. We use three sets of features (summarized in Table 4) to classify the performance of a player: *demographic* (I), *quality* (Q) and *topological* attributes (T). Demographic features include age, gender, province and region of birth. Quality features contain the number of winning matches, sets and points in a season as well as the initial and final position in the rank. At last, topological features refer to the network analysis metrics presented in the previous section, ranging from the role of the players in the network (i.e., centrality measures) to their structural connectivity (degree measures). The topological features were computed for each of the three networks $G^s_{matches}$, G^s_{sets} and G^s_{points}.

We conducted four experiments described in the following sections: (1) correlation analysis, (2) talent identification, (3) prediction of performance-driven classes, and (4) prediction of the player's position in the ranking.

Experiment 1: Correlation analysis

In this first part, we performed a correlation study between the groups of features described in the previous section and two different performance metrics: (a) the improvement within a season (i.e., the percentage of positions due to the points gained at the end of a season), and (b) the final ranking (the position in the ranking due to the achieved points at the end of a season). We computed metrics for both position and points (i.e. improvement position, improvement points, final-year position, and final-year points). The goal of this experiment is to assess the presence of a statistical dependence between success and intrinsic and extrinsic properties of an athlete. First, we analyzed the shape of the distribution for all the features in the model to check whether they are normal or skewed. We adopted two primary methods of assessing normality: numerical (relying on statistical tests) and graphical (relying on visual inspection). For the latter, the analysis of frequency distribution and QQplot gave us a first insight. In addition, we implemented three of the most used normality tests (see Razali and Wah (2011)): Shapiro (see Lilliefors (1967)), Kolmogorov-Smirnov (see Frank and Massey (1951)) and Kolmogorov-Smirnov corrected by Lilliefors (see Lilliefors (1967)). We then applied Pearson correlation in the presence of normality, otherwise, we computed Kendall rank and Spearman correlations.

Experiment 2: Talent identification

In this task we focused on the identification of *talents*, where a *talent* is defined as a young athlete with age between 8 and 14 that reaches the top 10% of the national rank at the end of the season. We addressed the problem using classification. In particular, given the demographic, quality, and topological features of an athlete from past seasons $[1 \ldots s-1]$, we aim at predicting a binary label (talent or not talent) at the end of the season s. The task implements a basic classification framework with Random Forest (an ensemble based on decision trees).

Experiment 3: Prediction of performance-driven classes

This task is a generalization of the talent identification problem where we focus on the entire population with the goal of predicting in which performance-driven class a player belongs to at the end of a season. To this extent, we first partition the global ranking in 5 classes that contain respectively the players in the 1st, 5th, 10th, 25th, and 50th percentile of the ranking. This subdivision reflects a discussion with domain experts and it represents the main classes of performers generally found in the table tennis community. For simplicity, we call them *stars*, *top players*, *good players*, *average players*, and *amateurs*. Since the male and female rankings are independent, we run two separated experiments by gender. Each experiment implements a multi-class classification task that takes in input the features described in Table 4. The experiment tests the features ability to predict the portion of the ranking a player is likely to fall into at the end of a season. We adopted the Python scikit-learn logistic regression classifier[3], with the limited-memory Broyden-Fletcher-Goldfarb-Shanno (lbfgs) solver and the cross-entropy loss method to deal with the presence of multiple classes. We adopt cross validation (5 folds) and compute accuracy, macro-averaged f-score, and cross-entropy loss[4] to evaluate the estimators performance.

Experiment 4: Prediction of a player's final position in the ranking

Finally, we focus on predicting the position of the players in the national ranking at the end of a season s given the demographic, quality, and topological sets of features in Table 4 estimated for season $s-1$. We perform this task as a regression task in which the regression function computes a score for each player such that sorting the players by the score placed them in a ranking position. The error of the regression function for each player is computed in terms of the absolute error between the true and the estimated position in the ranking. The accuracy of the predicted ranking was evaluated both in terms of correlation (between the athlete's score and her true ranking position) and in terms of the Kendall τ which is a standard measure for the agreement between two rankings (see Kendall (1938)). This task was challenging since the number of players (especially males) was large and some tests required large computational times. The computational complexity for the solvers that gave the best results is $O(t \cdot f \cdot n \cdot log(n))$ where t (10) is the number of independent random folds in which the training-set has been divided, f (35) is the number of features and n (7,220) is the number of players. For these reasons we reported only the results of the

[3] `http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html`

[4] The cross-entropy loss is the loss function used in (multinomial) logistic regression and its extensions with the perceptron of artificial neural networks. It is defined as the negative log-likelihood of the true labels given the predictions of a probabilistic classifier.

best performing tests, in the single setting that included all the sets of features.

Results

Correlation analysis

We report the results of the correlation analysis between the characteristics of athletes described in Table 4 and two performance indicators: (a) the improvement within a season, and (b) the final ranking. Since many features failed the normality test, we applied Spearman correlation coefficient in these cases. As regards features that correspond to topological properties of the networks, we first considered the directed networks, in which the edges are incoming for the winner player. Results on $G^s_{matches}$ are reported in Table 5. It includes only the most significant athletes' variables, ordered by their correlation with the performance indicators (correlation values with position are negative because lower the athletes' position better their performance). Results consistently indicate a high correlation of the athletes' performances with the topological properties, in particular, centrality measures: this suggests the positive role of the node position in the network to predict performance.

We notice that the variables that have the highest correlation results with the athlete's performance indicators are the topological network properties; in particular, indegree, which might be a proxy for performance by construction. Therefore we repeated the correlation analysis considering the undirected networks as well. The undirected networks, by loosing the edges directionality, discard any information on the performance of a player in a dyad, but maintain the connectivity patterns. The correlation analysis performs also a sort of validity check on the built networks. The results on the undirected version of $G^s_{matches}$ are reported in Table 6. To make easier the comparison with results on directed networks in Table 5 we maintain the list of variables in the same ordering. The results on undirected networks still show good correlation results between the athlete's performance and the topological properties of the network (consider in particular the network measures of eigenvector and PageRank centralities). We performed the same correlation analysis considering the networks G^s_{sets} and G^s_{points} as well, obtaining similar results.

These results give us some first hints on the positive role of the network, as elaborated in the following:

- the higher the centrality of an athlete, the higher her performance. The intuition behind this observation could be that a player who is central in a network played more times and had the opportunity to gain more experience by competing with a larger and more diverse set of adversaries.
- the topological features have a positive role in capturing indirect connections, allowing to compare even players that did not compete directly. This is the core point that we would like to stress here: the topological features of the connections of a node in a network, such as the centrality of the node and the centrality of its neighborhood, are indicators of the relevance of the node in the network. These are measures that should be kept even in a network with undirected edges and therefore discarding the effect provided by the direct information of the outcome of games.
- the occurrences of matches (captured by the measures of degree and degree centrality) and the number and type of adversaries a player competed with, is a significant information in itself in estimating the players' validity. The reason

	Final-year position	Final-year points	Improvement position	Improvement points
C-Eig	-0.74	0.88	-0.38	0.37
Pagerank	-0.71	0.80	-0.60	0.57
D-In	-0.58	0.64	-0.69	0.67
D-In (W)	-0.58	0.63	-0.70	0.68
W-Set	-0.57	0.61	-0.68	0.66
D-Set	-0.49	0.57	-0.51	0.45
W-Mat	-0.47	0.56	-0.64	0.63
Deg(W)	-0.46	0.46	-0.58	0.57
N-Mat	-0.46	0.46	-0.58	0.57
C-Deg	-0.45	0.46	-0.58	0.57
Deg	-0.45	0.46	-0.58	0.57
C-Clo	-0.35	0.23	-0.22	0.25
L-Set	-0.28	0.24	-0.39	0.39
C-Bet	-0.23	0.18	-0.51	0.52
C-Loa	-0.22	0.17	-0.50	0.52
D-Out(W)	-0.17	0.12	-0.27	0.29
AvND	-0.16	0.13	-0.23	0.24
D-Out	-0.15	0.12	-0.27	0.29
Age	-0.10	0.18	-0.22	0.23

Table 5. Spearman correlation in directed networks between the final ranking and the improvement in a season with several athletes and network features

could be that the matches occurrences might already have a positive effect on the players' performance: during the matches the athletes gain more confidence and a better mental attitude toward the competition.

In the following sections, we provide additional evidence on the positive, predictive effect of the network measures in complex tasks. In the discussion of the results of the remaining experiments, for the sake of brevity, we consider only the directed versions of the networks.

Talent identification

The young athletes with age between 8 and 14 in season 2011/2012 were 3,700 males and 679 females. Reaching the top 10% of the national ranking means to climb to more than the 821th position for males and the 81th for females. The results of the talent identification task are shown in Table 7, where the features are sorted by accuracy (from best to worst). The worse performing features were the personal information (I) and quality information (Q) while the feature information on network topology (T) worked better for both F-measure and accuracy (in particular as regards the topological information extracted from G^s_{sets} (feature T_{sets})).

Combining the 5 groups of features together will get the best result (86.74 F-measure; 96.42 accuracy). The best feature set contained the topological features (T) while Q and I features allowed the system to get the less promising predictive results. The results highlight the fact that topological information extracted from the graphs of table tennis matches are more predictive for talent identification than information on matches performance.

	Final-year position	Final-year points	Improvement position	Improvement points
C-Eig	-0.51	0.62	-0.39	0.38
Pagerank	-0.30	0.29	-0.52	0.51
W-Set	-0.59	0.63	-0.69	0.66
D-Set	-0.49	0.58	-0.52	0.44
W-Mat	-0.46	0.57	-0.59	0.56
Deg(W)	-0.43	0.45	-0.44	0.40
N-Mat	-0.47	0.47	-0.44	0.45
C-Deg	-0.42	0.45	-0.53	0.51
Deg	-0.42	0.45	-0.52	0.51
C-Clo	-0.42	0.48	-0.53	0.53
L-Set	-0.28	0.25	-0.45	0.44
C-Bet	-0.27	0.22	-0.44	0.46
C-Loa	-0.26	0.22	-0.23	0.19
AvND	-0.49	0.58	-0.65	0.62
Age	-0.03	0.13	-0.18	0.22

Table 6. Spearman correlation in undirected networks between the final ranking and the improvement in a season with several athletes and network features

Group of features	Macro F-Score	Accuracy
Q, I, $T_{matches}$, T_{sets}, T_{points}	86.74	96.41
$T_{matches}$, T_{sets}, T_{points}	86.53	96.36
Q, I, $T_{matches}$, T_{sets}	86.21	96.38
Q, I, T_{sets}, T_{points}	86.01	96.24
T_{sets}	86.02	96.18
$T_{matches}$, T_{points}	85.64	96.16
$T_{matches}$, T_{sets}	85.24	95.99
T_{points}	85.14	95.99
T_{points}, T_{sets}	84.93	95.72
T_{points}, T1	84.18	95.79
Q, I, $T_{matches}$	83.85	95.72
Q, I, T_{sets}	83.61	95.66
Q, I, T_{points}	82.73	95.45
$T_{matches}$	80.85	95.00
Q, I	81.16	94.92

Table 7. Talent identification task results. Random Forest obtained significantly etter results on both accuracy and F-measure using T type features. Features based on topological information improved talent prediction compared to personal and quality ones

Table 8. Results of the prediction of performance-driven classes for male players on G1

Features	Accuracy	Macro F-score	Cross-entropy loss
T+Q+I	0.913 (±0.018)	0.85 (±0.017)	-0.22 (±0.02)
T+Q	0.90 (±0.02)	0.85 (±0.03)	-0.25 (±0.04)
T	0.63 (±0.03)	0.39 (±0.05)	-0.92 (±0.08)
Q	0.83 (±0.02)	0.78 (±0.02)	-0.41 (±0.04)

Prediction of performance-driven classes

Results for both male and female players are summarized, respectively, in Table 8 and Table 9. It is worth mentioning that in the experiments we did not include any features (performance as well as topological) of the same season for which the prediction occurs but we considered only features from the previous years. Thus the validity of the predictions pretended to evaluate the possibility of a trainer to forecast the outcomes of the future season. In both the experiments, the combination of the three groups of features achieved consistently the highest value across all the performance metrics. Accuracy reaches 0.913 and 0.907 for males and females respectively: this shows the ability of our framework to model success as a categorical variable. Note that the features on matches performance have the strongest impact on prediction ($accuracy = 0.83$ if considered in isolation), followed by the topological (T) and then the personal information (I). As hypothesized, the contribution of the network-based features to the overall performance is significant (+8% over the quality features - Q), while the personal information such as age, city, province and region of birth did not produce a substantial contribution.

Table 9. Results of the prediction of performance-driven classes for female players on G1

Features	Accuracy	Macro F-score	Cross-entropy loss
T+Q+I	0.907 (±0.018)	0.80 (±0.16)	-0.50 (±0.52)
T+Q	0.88 (±0.06)	0.78 (±0.17)	-0.48 (±0.29)
T	0.70 (±0.10)	0.64 (±0.15)	-0.78 (±0.16)
Q	0.81 (±0.04)	0.74 (±0.15)	-0.43 (±0.15)

Prediction of a player's final position in the ranking

For sake of space we report only the results of SVM for regression that allowed us to obtain the best results among several methods (LASSO, LARS, Multi-Layer Perceptron, Elastic Net, Random Forest). We performed separate tests with sci-kit learn and Weka[5] for females and males whose results are shown in Figure 3 and Figure 4. In the diagrams the outliers are shown with bigger markers. For both, we employed a polynomial (linear) kernel and used 10-fold cross validation. It is evident the linear relationship between the regressors and the final score. In Table 10 we report the quantitative results.

The average absolute errors (MAE) in terms of positions in the ranking are still quite high (733 out of 6,553 (11%) and 86 out of 667 (12%) respectively in males and females). The errors decrease with the better ranking positions: 261 and 33 positions (resp. for males and females in the first two hundreds and one hundred positions of

[5]http://www.cs.waikato.ac.nz/ml/weka/index.html

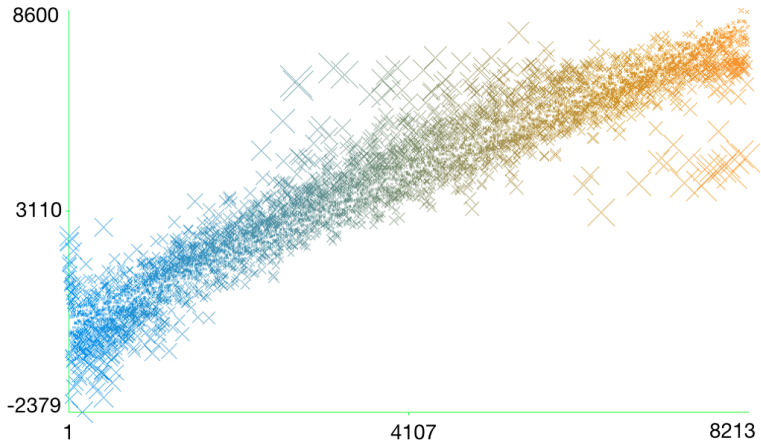

Figure 3. Regression by SVM on male athletes' final position in ranking

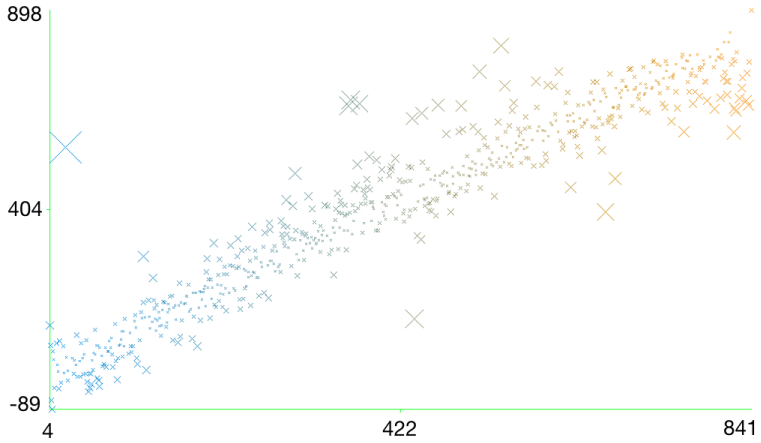

Figure 4. Regression by SVM on female athletes' final position in ranking

the ranking). As regards the best predictors[6] we mention the total number of points, the total number of sets in which the athletes won, some measures of networks that describe how strong the athletes connections with the others athletes are (eigenvector centrality, in-degree weight, betweenness centrality, degree weight, load centrality), the total number of matches and the number of sets in which the athletes lost (evidently they learn also from the lost sets). Finally, but with a much lower weight (half and even less), the geographic provenance of the athletes says something about its team: Guiglia (Emilia-Romagna), Canda (Veneto), Maratea (Basilicata), Reggello (Toscana), Casapulla (Campania), Pontelongo (Veneto).

Finally we evaluated the plausibility of the produced rankings and verified that the athletes in the positions at decreasing ranges of one hundred positions, showed some aggregated results on matches performance that are decreasing steadily as well as on the network topology measures.

[6]We report the results on males since the results for females are similar but the males population is almost ten times larger.

Season	gender	MAE	MAE%	MAE (@ first n)	corr.	Kendall τ
15/16	male	733 (out of 6,553)	11%	261 (@200)	0.97	0.87
15/16	female	86 (out of 667)	12%	29 (@50)	0.96	0.85

Table 10. Results of the regression for the prediction of the seasonal ranking

Conclusions

While the vast majority of studies on athletes performance focus their attention on the role of training, behavior, nutrition, or psychological attitudes, this paper shed light on the importance of the network of matches played amongst athletes in modeling success. Analyzing the entire set of Italian table tennis matches between 2011 and 2016, we were able to observe a positive correlation between several network centrality measures and the improvement in the final ranking. In other words, we quantitatively captured the relation between the quality, quantity, diversity of your opponents and the development of a player's career. Given this intuition, we used a machine learning framework to approach three different problems of increasing difficulty: namely, talent identification, prediction of a player's performance-driven class, and prediction of the final ranking. Results showed that combining profile, activity metrics and topological features gives consistently the best performance across tasks and seasons. These observations have potentially a strong impact on defining innovative methodologies to model performance in sports, and to encourage national federations to develop specific projects with a multidisciplinary and integrative perspective. For instance, the relevance of network metrics like centrality, could suggest to create more opportunities for players to participate in high-quality tournaments, i.e. encouraging table tennis federations to increment tournaments at national level. Similarly, societies might provide incentives to their most promising athletes to participate in such venues.

Acknowledgement

We are grateful for the collaboration of the Italian Table Tennis Federation (FITET - *www.fitet.org*).

References

Albert, R., & Barabási, A.-L. (2002, Jan). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, *74*, 47–97. Retrieved from `https://link.aps.org/doi/10.1103/RevModPhys.74.47`

Atkinson, G., & Reilly, T. (1996). Circadian variation in sports performance. *Sports Medicine*, *21*(4), 292–312. Retrieved from `http://dx.doi.org/10.2165/00007256-199621040-00005`

Bavelas, A. (1950). Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America*, *22*(6), 725-730. Retrieved from `https://doi.org/10.1121/1.1906679`

Ben-Naim, E., Redner, S., & Vazquez, F. (2007). Scaling in tournaments. *EPL (Europhysics Letters)*, *77*(3), 30005. Retrieved from `http://stacks.iop.org/0295-5075/77/i=3/a=30005`

Brin, S., & Page, L. (1998, April). The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. ISDN Syst.*, *30*(1-7), 107–117. Retrieved from `http://dx.doi.org/10.1016/S0169-7552(98)00110-X`

Cronin, J., & Sleivert, G. (2005). Challenges in understanding the influence of maximal power training on improving athletic performance. *Sports Medicine*, *35*(3), 213–234.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, *100*(3), 363.

Fast, A., & Jensen, D. (2006). The nfl coaching network: analysis of the social network among professional football coaches. *Growth*, *60*, 70.

Frank, J., & Massey, J. (1951). The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, *46*(253), 68-78. Retrieved from `http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1951.10500769`

Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the national academy of sciences*, *99*(12), 7821–7826.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *1/2*(318), 81–93.

Kondric, M., Sekulic, D., & Mandic, G. F. (2010). Substance use and misuse among slovenian table tennis players. *Substance Use & Misuse*, *45*(4), 543-553. Retrieved from `http://dx.doi.org/10.3109/10826080903452553` (PMID: 20141464)

Kovalchik, S. (2016, July). Latest reserach into sports technology. *The Australian Sports Analytics Conference*(6). Retrieved from `http://www.analyticsinsport.com/australia/`

Kusubori, S., Yoshida, K., & Sekiya, H. (2012). The functions of spin on shot trajectory in table tennis. In *Isbs-conference proceedings archive* (Vol. 1).

Lames, M., & McGarry, T. (2007). On the search for reliable performance indicators in game sports. *International Journal of Performance Analysis in Sport*, *7*(1), 62–79.

Lejeune, M., Decker, C., & Sanchez, X. (1994). Mental rehearsal in table tennis performance. *Perceptual and Motor Skills*, *79*(1), 627-641. Retrieved from `http://dx.doi.org/10.2466/pms.1994.79.1.627`

Lilliefors, H. W. (1967). On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of the American statistical Association*, *62*(318), 399–402.

McGarry, T., & Franks, I. M. (1994). A stochastic approach to predicting competition squash match-play. *Journal of Sports Sciences*, *12*(6), 573-584. Retrieved from `http://dx.doi.org/10.1080/02640419408732208` (PMID: 7853454)

Nevill, A., Atkinson, G., & Hughes, M. (2008). Twenty-five years of sport performance research in the journal of sports sciences. *Journal of Sports Sciences*, *26*(4), 413-426. Retrieved from `http://dx.doi.org/10.1080/02640410701714589` (PMID: 18228169)

Newman, M. (2010). *Networks*. Oxford University Press.

Newman, M. E. (2008). The mathematics of networks. *The new palgrave encyclopedia of economics*, *2*(2008), 1–12.

Newman, M. E. J. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 404-409. Retrieved from `http://www.jstor.org/stable/3054694`

Newman, M. E. J. (2002, Oct). Assortative mixing in networks. *Phys. Rev. Lett.*, *89*, 208701. Retrieved from `https://link.aps.org/doi/10.1103/PhysRevLett.89.208701`

Nicholls, A. R., & Polman, R. C. J. (2007). Coping in sport: A systematic review. *Journal of Sports Sciences*, *25*(1), 11-31. Retrieved from `http://dx.doi.org/10.1080/02640410600630654` (PMID: 17127578)

Onody, R. N., & de Castro, P. A. (2004). Complex network study of brazilian soccer players. *Physical Review E*, *70*(3), 037103.

Park, J., & Newman, M. E. (2005). A network-based ranking system for us college football. *Journal of Statistical Mechanics: Theory and Experiment*, *2005*(10), P10014.

Pion, J., Hohmann, A., Liu, T., Lenoir, M., & Segers, V. (2017). Predictive models reduce talent development costs in female gymnastics. *Journal of Sports Sciences*, *35*(8), 806-811. Retrieved from `http://dx.doi.org/10.1080/02640414.2016.1192669` (PMID: 27267568)

Raab, M., Masters, R. S., & Maxwell, J. P. (2005). Improving the 'how' and 'what' decisions of elite table tennis players. *Human Movement Science*, *24*(3), 326 - 344. Retrieved from

http://www.sciencedirect.com/science/article/pii/S0167945705000394 (Motor control and learning: A contribution from the European School of Sports Science)

Radicchi, F. (2011, 02). Who is the best player ever? a complex network analysis of the history of professional tennis. *PLOS ONE*, *6*(2), 1-7. Retrieved from https://doi.org/10.1371/journal.pone.0017249

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, *2*(1), 21–33.

Sulis, E., Meo, R., Schifanella, R., & Arisi, A. (2017). A network analysis of table tennis matches played in italy. *Proceedings Book of the 15th ITTF Sports Science Congress, Düsseldorf, May 27th - 28th 2017*, 58-63.

Walling, M. D., Duda, J. L., & Chi, L. (1993). The perceived motivational climate in sport questionnaire: Construct and predictive validity. *Journal of Sport and Exercise Psychology*, *15*(2), 172–183.

Wasserman, S., & Faust, K. (1995). *Social network analysis.* Cambridge University Press.

Woodman, T., & Hardy, L. (2003). The relative impact of cognitive anxiety and self-confidence upon sport performance: A meta-analysis. *Journal of sports sciences*, *21*(6), 443–457.

Woods, C. T., Raynor, A. J., Bruce, L., McDonald, Z., & Robertson, S. (2016). The application of a multi-dimensional assessment approach to talent identification in australian football. *Journal of Sports Sciences*, *34*(14), 1340-1345. Retrieved from http://dx.doi.org/10.1080/02640414.2016.1142668 (PMID: 26862858)