Optimal coaching: How the tools of artificial intelligence and mathematics can help

Mathieu Rosenbaum

École Polytechnique-ALYX

28 june 2023

Table of contents

Our research group : A team of quants

We have a quantitative finance background. In finance our modeling approach goes as follows :

Original topic of interest : Quantitative Finance

Step 1 : Data science.

Analyze market data to understand asset dynamics and their correlations.

- *We estimate from past data the trends and volatilities of Google, Apple, Boeing, BNP-Paribas...*
- *From historical data, we measure the correlations between these assets.*

Our research group : A team of quants

Original topic of interest : Quantitative Finance

Step 2 : Probabilistic modeling/AI.

Based on our statistical analysis, we build market simulators to be able to make quantitative forecasts about the results of our own strategies, notably in term of risk.

- *We use probabilistic tools enabling us to mimic the future behaviors of Google, Apple, Boeing, BNP-Paribas.*
- *We must ensure that in the future behaviors of these assets, the statistical properties established above remain satisfied.*

Our research group : A team of quants

Original topic of interest : Quantitative Finance

Step 3 : Optimization (Stochastic control).

Using our probabilistic modeling, we can optimize our strategies by finding the best combinations of assets and transaction times to optimize our profit and loss.

- *I have a given amount of money that I can invest in Google, Apple, Boeing and BNP-Paribas.*
- *At starting time of the investment, how do I split my money between these assets according to my performance criterion.*
- *How and when to modify it through time if I am unhappy with the performance ?*
- *How relevant were my modifications ?*

Our research team

From finance to football

We adopt the same paradigm for football modeling developing a tool where :

- Assets→Players
- Markets→Games
- Portfolio strategy→ Selected players, tactic, change of tactic/substitutions
- Profit and loss→Score
- Asset Manager→Coach

The model does not replace the experienced asset manager/coach, but help him in his decision process. We present this tool in this talk.

Table of contents

Simulating games

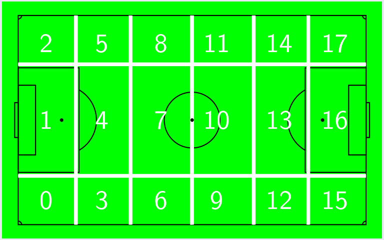
Aim

- Simulating a game between two given teams.
- Goal : Forecasting relevant quantities : score, time with the ball of each team, names of the strikers...
- This will enable us to optimize starting team and coach strategy.
- Simulator can be used even if players do not usually play together (interesting for national teams, scouting, see later).
- We use the following 18 zones.

Framework

Pitch decomposition

- The field is split into 18 zones.



Description of the model

The model

Each player i_1 in zone $j_1 \in \{0, \ldots, 17\}$ can :

- Give the ball to a partner i_2 in zone j_2.
- Lose the ball to a player of the other team i_2 in zone j_2.
- Score.

The probability of these events depends on :

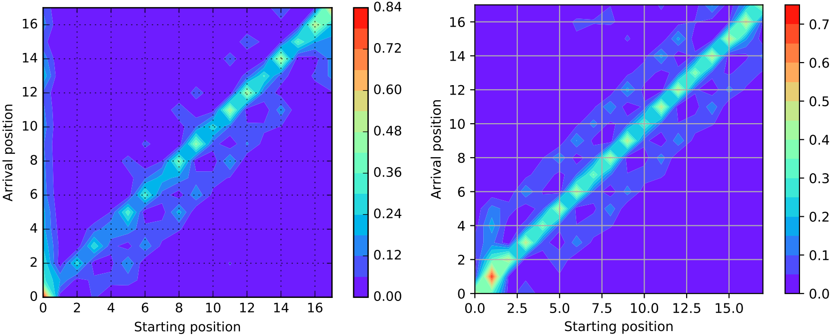
- State of the game : players on the field, players positions, coach strategy, score, time...
- The ability of player i_1 to transfer the ball in zone j_2.
- The probability of presence of player i_2 in position j_2.
- The ability of the partner (resp. player from the other team) i_2 to receive (resp. to intercept) the ball in zone j_2.
- The ability of the other team to intercept the ball in j_2.

Calibration of the model

Artificial intelligence/machine learning

- When a given player has the ball at a given location, his next action and its success or failure are decided in a random fashion, yet depending on the player abilities and its environment through a mathematical model.
- The key explanatory variables for the success or failure are selected through a thorough statistical (big data) study, corresponding to Step 1 above.
- Model calibrated on Opta F24 event data.
- Calibration method : Artificial intelligence algorithms/machine learning.

Transition probabilities in data (left, Ligue 1 2016/17) and in the model for a PSG-Lyon game



Transition probabilities for E. Cavani starting from Zone 14, in data (left, Ligue 1 2016/17) and in the model

0	0	0.023	0.1	0.38	0.068
0	0	0	0.034	0.15	0.034
0	0	0	0.011	0.15	0.057

0	0	0.015	0.12	0.44	0.11
0	0	0	0.029	0.13	0.035
0	0	0	0.007	0.081	0.033

Final ranking in simulated return phase of Ligue 1 2016/17 (calibration on first half season, 100 simulations)

Team	Average ranking	Standard deviation
Monaco	2.2	1.2
PSG	4.0	2.9
Nice	4.1	2.7
Bordeaux	4.6	2.2
Marseille	6.3	2.0
Angers	7.5	4.9
Lyon	8.7	3.9
Nantes	9.9	4.3
Lille	9.9	4.8
Dijon	11.0	4.5
Bastia	11.6	4.2
St-Etienne	12.6	4.7
Rennes	12.9	3.7
Montpellier	13.6	3.3
Guingamp	14.1	5.1
Lorient	14.5	3.7
Nancy	14.7	2.0
Metz	14.7	3.0
Toulouse	15.7	5.7

Side application : Scouting

Finding the most suitable new player

- Based on our approach, we can find the players who resemble a given player the most. We can quantify how close the players are.
- We consider the variables found relevant in our simulation device.
- We can also look for a player resembling a fake player with given characteristics.
- Example : the 10 players from the five main championships in 2017/18 that are the closest to Thiago Motta 2016/17.

Side application : Scouting

The Thiago Motta case

Closest players to Thiago Motta (in order) :

- Julian Weigl (Borussia Dortmund)
- Maxime Gonalons (AS Roma)
- Javi Martinez (Bayern Munich)
- Lucas Tousart (Olympique Lyonnais)
- Wylan Cyprien (OGC Nice)
- Mario Lemina (Southampton)
- Jean-Michael Seri (OGC Nice)
- Fernandinho (Manchester City)
- Stanislav Lobotka (Celta Vigo)
- Rodrigo Bentencur (Juventus Turin)

Table of contents

Optimization framework

Stochastic control before the game

- Before starting the game, we define a performance criterion (difference of goals, number of points at the end of the game...).
- We have a stochastic dynamic for the game (simulator). The dynamic will depend on our choices : composition and tactic.
- In this setting we can use stochastic control theory to find the optimal decision (composition and tactic) to take before the game and to assess the value of any decision.

Optimization framework

Stochastic control before and during the game

- We call value function $v(t, u)$ the expected score at the end of the game if we act optimally if there are t minutes remaining and the situation of the game is u (composition, score, etc.).
- Optimal decision before the game leads to $v(T, u_0)$.
- Then, during the game, we can monitor in continuous time the value function, optimal decisions and the value associated to other decisions.
- We deduce optimal substitutions/tactic changes and optimal times to do them.

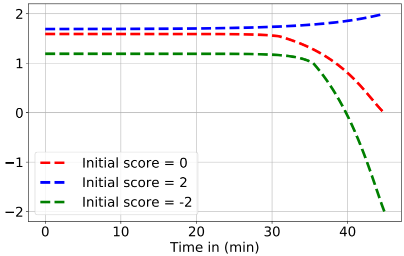
Example

Behavior of the value function

- Round of 16, Champions league 2017/2018 : PSG-Real Madrid.
- Substitution : Bale replaces Benzema at 76th minutes.
- Compare to the case with no change, jump in the value function of 4%.
- $v(76, \text{Bale}) = 1.04 * v(76, \text{Benzema})$.

Evolution of the value function

$V(t, u_t)$ when s moves



$V(t, u_t)$ when m moves

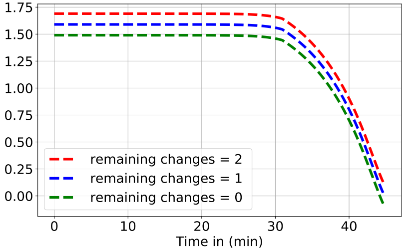


FIGURE – Evolution of the value function in a second half-period depending on score and number of remaining substitutions (PSG-Lyon).

Ex. of optimal decision : French team (pre world-cup data)

Against Brasil

- Lloris/Sidibé-**Varane-Koscielny**-Mendy/Tolisso-Kanté-Matuidi/Mbappé-Griezmann-Lemar.

Against Germany or Spain

- Lloris/Sidibé-**Koscielny-Umtiti**-Mendy/Tolisso-Kanté-Matuidi/Mbappé-Griezmann-Lemar.

Against Portugal

- Lloris/Sidibé-**Varane-Koscielny**-Mendy/Mbappé-Kanté-Matuidi-Lemar/Griezmann-**Giroud**.

Against Uruguay

- Lloris/Sidibé-**Varane-Umtiti**-Mendy/**Pogba**-Kanté-Matuidi/**Dembélé-Mbappé-Lemar**.

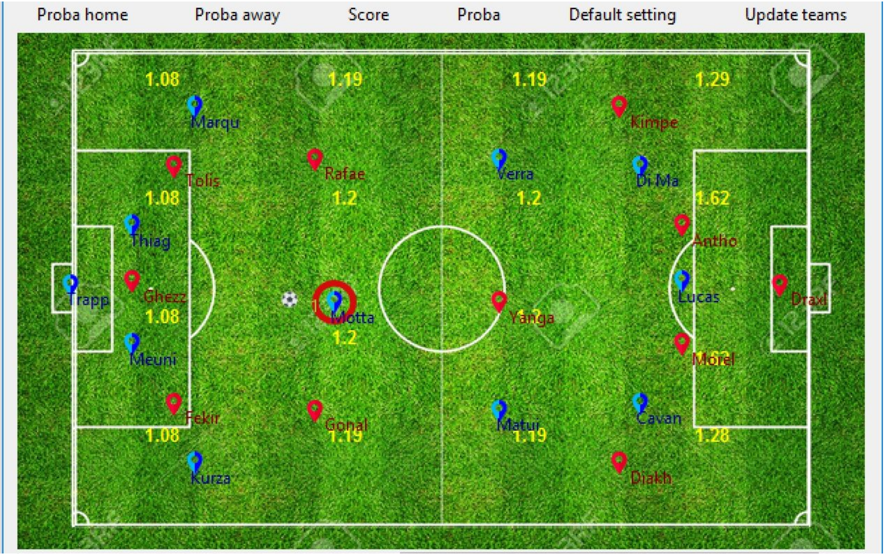
Quantifying any past decision

Did the player make the right choice ?

When the player has the ball at a given point, and for a given configuration of the other players (both teams), based on our model we can :

- Predict what the player will do : give a probability to all the possible next locations for the ball.
- For each possible event, probability of success or failure.
- Quantify the value of each possible action (if succeeded).
- Deduce the optimal action (expected utility).

Value function at next time if attempt successful



Probability success of a pass

Screenshot of Lyon vs Paris Saint-Germain

- One obviously needs to discount the previous values !
- Probability of successful pass in a zone (yellow).
- Probability of missed pass in a zone (white).

Probability success of a pass

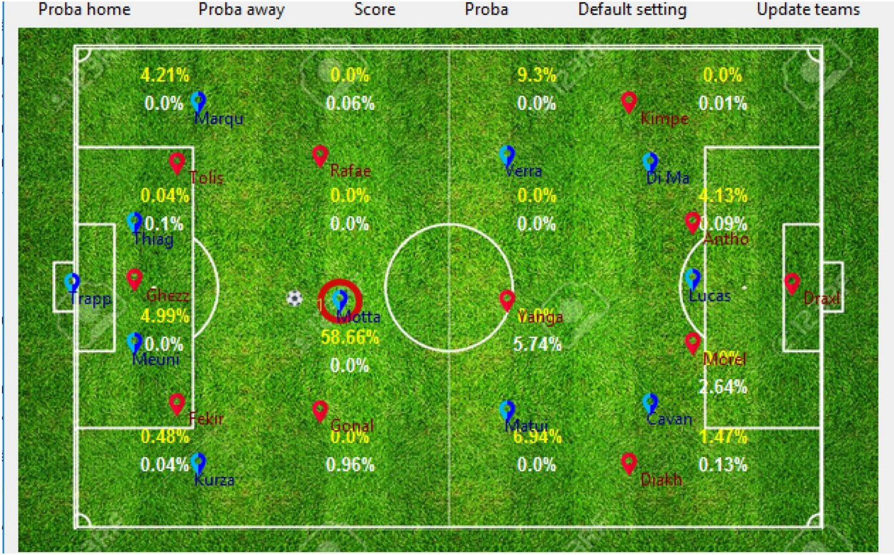


Table of contents

Definition

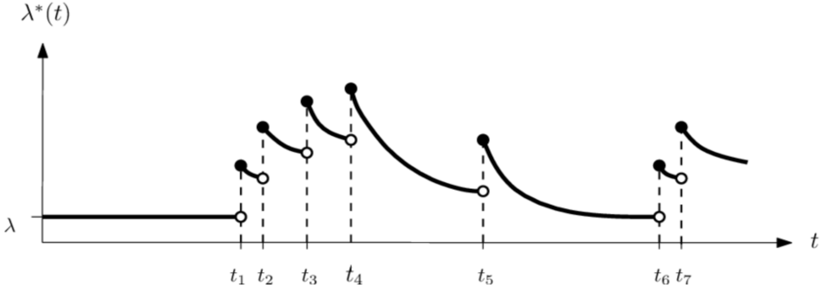
Hawkes process

- A Hawkes process $(N_t)_{t\geq 0}$ is a self-exciting point process, whose intensity at time t, denoted by λ_t, is of the form

$$\lambda_t = \mu + \sum_{0<J_i<t} \phi(t - J_i) = \mu + \int_{(0,t)} \phi(t-s)dN_s,$$

  where μ is a positive real number, ϕ a regression kernel and the J_i are the points of the process before time t.
- These processes have been introduced in 1971 by Hawkes in the purpose of modeling earthquakes and their aftershocks.
- Used in neurosciences, network analysis, criminology, finance...

Example of Hawkes process intensity



Popularity of Hawkes processes in finance

Two main reasons for the popularity of Hawkes processes

- These processes represent a very natural and tractable extension of Poisson processes. In fact, comparing point processes and conventional time series, Poisson processes are often viewed as the counterpart of iid random variables whereas Hawkes processes play the role of autoregressive processes.
- Another explanation for the appeal of Hawkes processes is that it is often easy to give a convincing interpretation to such modelling. To do so, the branching structure of Hawkes processes is quite helpful.

Hawkes processes as a population model

Poisson cluster representation

- Under the assumption $\|\phi\|_1 < 1$, where $\|\phi\|_1$ denotes the L^1 norm of ϕ, Hawkes processes can be represented as a population process where migrants arrive according to a Poisson process with parameter μ.
- Then each migrant gives birth to children according to a non homogeneous Poisson process with intensity function ϕ, these children also giving birth to children according to the same non homogeneous Poisson process, see Hawkes (74).
- Now consider for example the classical case of buy (or sell) market orders. Then migrants can be seen as exogenous orders whereas children are viewed as orders triggered by other orders.

Table of contents

Event based data

Event based football data

- A time-coded feed that lists all events with the ball within a game with a player, team, event type and timestamp for each action.

Counting process

A 12-dimensional counting process is constructed as follows :

- Every time a player touches the ball, there is a jump in his assigned dimension $d \in \{1, \ldots, 11\}$ at the corresponding timestamp.
- Every time the ball is in the opponent danger area, there is a jump in the twelfth dimension at the corresponding timestamp.
- Once a danger state is triggered, no event are recorded until the ball goes out of the danger area $(+\epsilon)$.
- Once the ball is lost, no event is recorded until the ball is won again.

Modelling with Hawkes process

Modelling and estimation

- We fit a 12-dimensional exponential Hawkes process to the jump data through Maximum Likelihood Estimation.

$$\lambda_i(t) = \mu_i + \sum_{j=1}^{d} \int_0^t \phi_{i,j}(t-s) N_j(ds).$$

  where $\phi_{i,j}(s) = \alpha_{i,j} e^{-\beta_i s}$.
- The goal is to detect correlations between the event times related to each player and the danger state.
- The estimation is complex in large dimensions as the likelihood function is not convex in β. This is why we reduce the number of parameters and consider the rate of decay of the kernel is the same for each player. $\beta_{i,j} = \beta_i$ for all i, j.
- We use the algorithm of Bonnet *et al.* (2022).

Estimated parameters

- The integrated kernel $K_{i,j} = \int_0^\infty \phi_{i,j}$ can be interpreted as the expected number of touches of player i **directly** generated by player j touch. It can be estimated directly from the estimated parameters :

$$\hat{K}_{i,j} = \hat{\alpha}_{i,j} / \hat{\beta}_i.$$

- It is also relevant to consider the interactions between two states across multiple steps (generations). This is the case for defenders that rarely generate Danger directly but contribute to Danger creation by passing the ball to advanced positions.
- $M_{i,j}$ represents the expected total number of touches of the player i generated by player j directly but also indirectly through other players.

$$M = K + K^2 + K^3 + \cdots = K\left(I - K\right)^{-1}.$$

where $K = \int_0^\infty \phi$.

Estimated interactions

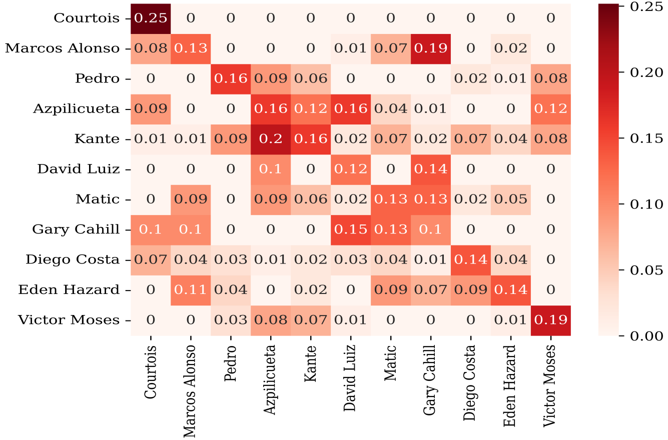


FIGURE – Graph of estimated interactions $\hat{K}_{i,j}$ between players for Chelsea 2016/2017.

Danger generated

Player Name	Danger directly generated	Total danger generated
Eden Hazard	0.22	0.28
Pedro	0.12	0.17
Victor Moses	0.08	0.15
Kante	0.06	0.13
Diego Costa	0.07	0.13
Marcos Alonso	0.04	0.10
Azpilicueta	0.02	0.10
Matic	0.03	0.10
Cahill	0	0.07
Courtois	0	0.05
David Luiz	0	0.04

TABLE – Danger directly and indirectly generated by one player touch. Direct danger is represented by $\hat{K}_{\text{danger},player}$ and Total danger by $\hat{M}_{\text{danger},player}$.

Some observations

Remarks

- Factoring in the indirect contribution in danger creation is important for defenders and midfielders.
- Kante is responsible for as many intrusions to the danger area per touch as the striker Diego Costa, while having more touches per game.
- The contribution of the central defender David Luiz in danger creation is minimal. This is not surprising as the flat 3-4-3 system relies heavily on the wings. David Luiz naturally passes the ball to either Cahill or Azpilicueta in build-up to spread the play.

A better visualization scheme

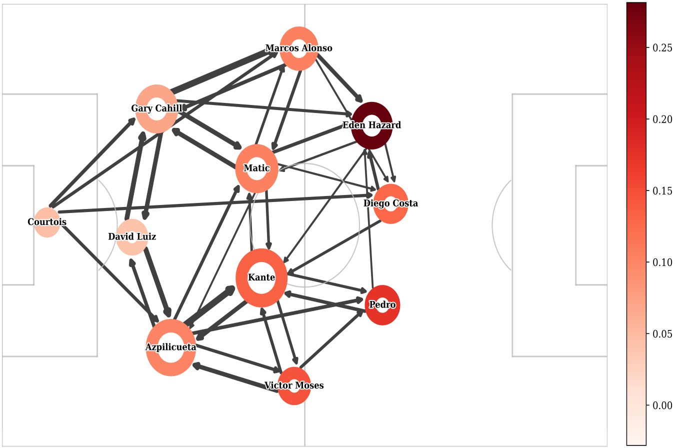


FIGURE – Graph of estimated interactions between players and with Danger state for Chelsea 2016/2017. The color of the circles represents the total danger created by the player. The size of arrows represents the parameters $\hat{K}_{player_1,player_2}$.

Comments

More remarks

- Azpilicueta, Kante and Victor Moses witness a considerable increase in Danger creation when considering the total contribution rather than the direct one. The right side of Chelsea combines a lot for danger generation and should be disrupted from the root.
- The left side relies a lot more on the huge offensive output of Eden Hazard. The links Marcos Alonso/Matic $\rightarrow$ Hazard should be controlled.
- Goalkeeper Courtois is successful in targeting Marcos Alonso and Diego Costa when playing long balls.