



ÉCOLE CENTRALE LYON

PAR
MÉCASPORTCo

Rapport intermédiaire

Élèves :

Nathan RIVIÈRE
Gabin ROLLAND

Enseignants :

Romain VUILLEMOT
Wouter BOS
Vincent FRIDRICI

9 janvier 2019

Table des matières

Introduction	2
1 Définition du sujet	3
1.1 Contexte	3
1.2 État de l'art	3
1.2.1 Autour de l'occupation d'espace	3
1.2.2 Possibilité de passes et analyse des tirs	5
1.2.3 Liens entre sport et sciences	6
1.2.4 Interface visuelle	6
1.3 Source de données disponibles	7
1.4 Objectifs	7
2 Résultats intermédiaires	9
2.1 Introduction	9
2.2 Démarche générale	9
2.3 Démarche détaillée	9
2.3.1 Diagrammes de Voronoï	9
2.3.2 Évaluation de l'occupation d'espace	10
2.3.3 Prise en compte de l'inertie	11
2.4 Calcul des aires des zones contrôlées	13
2.5 Analyse des résultats	14
2.5.1 Influence de l'inertie sur les espaces de Voronoï	14
2.5.2 Le lien avec la prédiction de passe	16
2.5.3 Remarques diverses	18
2.6 Espace partagé	19
2.7 Travaux à venir	19
3 Conclusion	21
Conclusion	21
Annexe A	22
Annexe B	23

Introduction

Aujourd'hui la technologie intervient de plus en plus dans le sport. En effet les statistiques des joueurs sont de plus en plus précises et de plus en plus nombreuses. Les clubs emploient des personnes pour acquérir des données et les exploiter. Cependant ces recherches ne sont pas disponibles. C'est dans ce contexte que s'inscrit le PAr MécaSportCo dont le but est d'analyser des données sportives pour trouver un ou plusieurs critères qui pourrait être utiles aux experts.

Ce livrable intermédiaire a pour but de présenter les objectifs de notre travail de recherche, l'état de l'art ainsi que les résultats obtenus. Il permet ainsi de faire un point sur le travail effectué durant cette première phase du projet. Une première partie détaille le contexte, ce qui a été fait dans ce domaine et rappelle les objectifs de notre travail. Puis, une seconde partie présente nos premiers résultats et le travail à venir.

1 Définition du sujet

1.1 Contexte

L'analyse de données devient essentielle dans le sport et est de plus en plus développée. Le principe est de collecter diverses données et de les analyser afin d'effectuer des prédictions ou d'améliorer les performances à la fois collectives et individuelles des joueurs. De nombreuses startups se créent et proposent leurs services aux professionnels et amateurs. Cette pratique est utilisée par des sportifs de haut niveau comme par exemple l'équipe de football d'Allemagne qui a eu recours à toutes les données disponibles sur leurs matchs pour se préparer à la coupe du monde de football 2014 qu'elle a par ailleurs remportée. On comprend donc ici l'importance de s'intéresser à cette activité, d'autant plus que la mise à disposition des données étant récente, tout n'a pas été exploré et un large champ de recherche est possible.

De plus l'exploitation de ces données d'un point de vue autre que sportif peut s'avérer pertinent comme c'est le cas dans l'article [1] qui illustre l'existence de certaines corrélations entre les mouvements de joueurs de football sur un terrain et des particules sur ce même terrain. C'est dans ce cadre qu'a été créé le PAr MécaSportCo lancé en 2017 afin de caractériser les mouvements de sportifs tout en cherchant à vérifier ces corrélations. Un premier travail a été effectué par Marc Louis Mattis et Alfonso García Hernández dont les principaux résultats sont mentionnés dans la section 1.2. Ce projet est ainsi continué cette année avec pour objectif de progresser dans l'exploration des différentes caractérisations possibles des comportements des sportifs à partir de l'exploitation de différentes données ainsi que dans le développement d'une interface visuelle permettant de communiquer ces données.

1.2 État de l'art

1.2.1 Autour de l'occupation d'espace

Les données sportives les plus simples à acquérir sont les positions des joueurs et du ballon. Ainsi étudier l'occupation de l'espace par les sportifs dans les sports collectifs constitue une analyse développée par plusieurs chercheurs. Les principales clés de l'occupation d'espace sont : l'interaction entre joueurs, les zones d'influences des joueurs, les options de passe, les espaces libres.

Les espaces de Voronoï Plusieurs articles mentionnent les espaces de Voronoï afin de caractériser l'occupation du terrain par les différents joueurs. Ainsi D. Cervone [2] découpe le terrain de basket en cellules grâce aux diagrammes de Voronoï et chacune de ces cellules est associée à un joueur. Au sein de chacun de ces espaces l'influence d'un joueur est inversement proportionnelle à la distance à laquelle il se situe.

Marc, dont le travail a précédé le notre, a effectué plusieurs études autour des espaces de Voronoï. La première étude concerne d'éventuels liens entre les espaces de Voronoï (qui caractérisent la zone d'influence) des joueurs et leur vitesse, la possession de balle et le moment où ils marquent un point. Ainsi il y aurait une corrélation entre la surface de ces espaces et la vitesse des joueurs au foot mais pas au basket. En revanche en regardant l'évolution de la surface pour les joueurs en possession du ballon on n'observe a priori

pas de lien exploitable au foot. Enfin lorsqu'on observe l'espace de Voronoï d'une équipe de basket juste avant qu'elle marque on constate simplement que la surface moyenne augmente puis diminue au moment du tir.

Zone d'influence des joueurs Pour caractériser l'influence d'un joueur sur une partie du terrain il est possible d'utiliser d'autres outils que les espaces de Voronoï.

Par exemple M. Stein et al. [3] développent un modèle permettant de caractériser la zone d'influence des joueurs de football en prenant en compte leur vitesse. Chaque joueur possède donc une zone qu'il contrôle autour de lui, plus sa vitesse est grande plus il occupe une zone étirée (dans la direction de sa vitesse) mais étroite.

A partir de l'influence des joueurs il établit ensuite des zones d'interactions comme l'intersection des zones d'influence de chaque joueur. Cependant dans cette démarche on constate qu'un espace important du terrain n'est pas occupé : ce sont les espaces libres. Ces espaces sont donc associés au joueur qui a une plus grande probabilité d'arriver en premier en prenant en compte sa vitesse, sa direction et sa distance.

D'une manière différente, L. Bornn [4] caractérise l'influence d'un footballeur grâce à une densité de probabilité d'une loi gaussienne à deux variables qui prend en compte la vitesse du joueur (norme et direction). À partir de ce critère, L. Bornn caractérise le contrôle d'une zone par une équipe en comparant l'addition des zones d'influences de chaque joueur par rapport à la somme de celles de l'équipe adverse. Ensuite, L. Bornn cherche à attribuer de la valeur aux zones du terrain selon les critères suivant : position par rapport à la balle, au but et des autres joueurs. Afin de déterminer une fonction caractéristique de la valeur d'une zone, il étudie d'abord la somme d'influence de la défense d'une zone en fonction de la position de la balle. Pour cela, il utilise un feed forward neural network. Une fois la valeur d'une zone déterminée à partir de la position de la balle et des défenseurs il la normalise par la valeur d'une zone vue par rapport à la position du but adverse.

L. Bornn avait déjà exploité cette idée au basket dans [2] en remarquant si un joueur i donne la balle à un joueur j c'est que celui-ci occupe une position mieux placée. À l'aide des données de position des joueurs et de passes, il avait associé une valeur à chaque espace du court. Il remarquait notamment que les zones importantes du terrain pouvaient varier d'une équipe à une autre.

L. Bornn dans [4], à partir de ses études sur les zones d'influence, le contrôle des zones et leur valeur se concentre finalement sur le sens des déplacements des joueurs avec deux critères : SOG (space occupation gain) et SGG (space generation gain).

Notons que dans [5], G. Andrienko et al. utilise également toutes ces idées afin de déterminer un moyen de mesurer la pression exercée sur joueur de foot par les autres joueurs (et le terrain).

Étude des trajectoires Pour étudier l'occupation d'espace, il est possible de se concentrer sur les trajectoires des joueurs et du ballon. C'est ce que font C. Kang et al. dans [6] en prenant en compte, à partir d'un instant t , la position que la balle et les joueurs pourraient avoir à l'instant $t+dt$. En prenant en compte les espaces possibles d'occupation des objets (joueurs et balles) et en regardant leurs intersections/exclusions ils définissent plusieurs critères : *possible region* (région qu'un objet peut potentiellement atteindre en un temps

donné, à partir d'un temps donné t), *catchable region*, *safe region*, *competition region*. A partir de ces critères plusieurs mesures peuvent être calculées comme l'aire moyenne des *safe region* des joueurs, ce qui permet de caractériser leur capacité à se démarquer.

Étude du centre de masse Le centre de masse d'une équipe est une donnée basique qui peut permettre de caractériser les mouvements collectifs.

Marc a travaillé sur ce critère en observant tout d'abord les comportements du centre de masse de deux équipes opposées au foot et au basket. Cette étude a permis de révéler une symétrie entre le mouvement des deux équipes (quand une avance l'autre recule) en particulier au foot. Ceci est une information importante car elle peut constituer un premier critère pour déterminer si une équipe attaque ou défend (positionnement du centre de masse par rapport à l'équipe adverse + vitesse relative).

De plus lorsqu'on compare la position des footballeurs relativement au centre de masse de leur équipe, Marc a constaté une organisation type de l'équipe ce qui peut permettre de déterminer si une équipe travaille bien en bloc ou pas et d'identifier un éventuel point faible dans le positionnement. Toutefois ceci n'est vrai que pour le foot car au basket et au rugby cette étude ne donne pas de résultats particuliers sur un match entier. Ceci est probablement dû au fait qu'il y a beaucoup de phases de jeu différentes et de changement de positions dans ces sports. Il faudrait donc réduire l'étude à certaine phase de jeu et non au match entier. Ce constat nous montre donc qu'il est intéressant de transposer les méthodes d'études sur plusieurs sports.

Le centre de masse peut également être étudié pour un groupe de joueurs impliqués dans une phase de jeu particulière au football plutôt que de regarder des valeurs moyennes pour toute une équipe. Ainsi, B. Santos et al. [7] apportent un moyen visuel de description d'une phase de jeu basé sur trois critères collectifs en lien avec le centre de masse : position du centre de masse, dispersion de l'équipe (position moyenne des joueurs par rapport au CDM, vitesse de dispersion ou de resserrement, distance occupée sur une largeur du terrain) et la synchronisation de l'équipe. Les critères individuels sont également utilisés mais de manière secondaire. Le principe de la description est le suivant : l'expert qui utilise l'outil sélectionne une action, pour cette action il choisit les joueurs qu'il qualifie d'important à ce moment et à partir de ces données l'interface met en évidence de manière visuelle les critères précédents. C'est donc un outil d'aide à la description d'action.

1.2.2 Possibilité de passes et analyse des tirs

L'étude des passes est particulièrement importante dans les sports collectifs. Plusieurs chercheurs travaillent autour de cette thématique.

Dans [8] Yue et al. détaillent la démarche qui a permis d'obtenir une interface visuelle de prédiction de passes au basket. Cette interface permet à l'utilisateur de placer lui-même les joueurs et le ballon sur le terrain et à partir de ceci on observe à l'aide de traits plus ou moins épais les probabilités de passe et de tir du joueur en possession de la balle. Le principe du calcul repose sur du machine learning expérimenté sur les données 2012-2013 de SportsVu.

M. Stein et al. [9] reprennent les modèles autour de l'occupation d'espace de l'article [3] afin de proposer une procédure d'analyse d'une action en terme de passe. On constate ici que l'étude des passes est étroitement liée à l'étude de l'occupation d'espace. Dans un

premier temps, il sélectionne les possibilité de passes ayant un intérêt grâce aux critères précédents. Ensuite la procédure consiste à regarder si la passe se fait dans le camp adverse ou non. Puis si la passe est dans le camp adverse il s'agit de prendre en compte l'influence des joueurs adverses. Enfin la dernière phase consiste à analyser ce que les défenseurs auraient pu faire pour empêcher la passe. Une procédure similaire est mise en place pour décrire les passes dans les espaces libres.

Il est également possible de se concentrer sur l'analyse des tirs comme le font Goldsberry [10] et Cervone [11] au basket. Goldsberry a développé un outil d'analyse pour la NBA qui permet de comparer visuellement les performances aux shoots des différents joueurs de basket. Il permet également de mettre en évidence les différents zones stratégiques pour les shoots sur un terrain de Basket. Dans l'article [11] Cervone et al. présentent une nouvelle métrique qui permet de caractériser une situation donnée en terme de possibilités de points marqués suite à cette action. Cette quantité est appelée EPV : *Expected Possession Value*. Le calcul de l'EPV se fait de manière statistique en prenant en compte une quantité finie de possibilités d'action pour le joueur en possession de la balle : dribbles, passes, shoots...

1.2.3 Liens entre sport et sciences

Corrélation entre particules fluides et sportifs L'étude des écoulements turbulents des particules fluides montre que la géométrie de l'espace dans lequel elles évoluent influe sur le changement de direction moyen. Il se trouve que la fonction de densité des changements de direction des joueurs de football sur un terrain de foot révèle une forme commune à celles de particules fluides confinées en 2D. Ainsi le comportement anisotrope et a priori non aléatoire des joueurs de foot n'a pas d'influence et seule la géométrie du domaine de limitation a un impact sur les changements de directions moyens [1].

Recherche de la course parfaite Aftalion et al. [12] proposent des modèles mathématiques permettant de simuler une course parfaite en athlétisme et d'étudier l'influence des entraînements sur les performances des athlètes.

1.2.4 Interface visuelle

La visualisation de données sportives peut prendre de nombreuses formes différentes. On peut représenter sous forme de graphique les statistiques de joueurs pour comparer leur performance sur différents critères. cette représentation ne fait que restituer les informations qu'apporte les données brutes. On peut présenter ces données de manière plus concrètes en affichant des graphiques superposés au terrain de sport. Enfin il est également possible de visualiser les données directement sur les vidéos des phases de jeu concernées afin de faire un lien concret entre données et phase de jeu [13].

SoccerStories [14] SoccerStories est une interface permettant de visualiser certaines phases de match de football. Cette interface peut s'avérer très utile en prenant en compte le fait que les données sous leur forme pure parlent peu. L'interviews d'experts confirment cette remarque puisque celles-ci révèlent que dans l'analyse du football il ne faut pas oublier le côté visuel et sélection de phases de jeu qui permettent de raconter une histoire. Nous pouvons puiser plusieurs inspirations de cette interface notamment les différents

critères et curseurs présentés en les adaptant à nos propres critères ainsi que les différentes démarchent présentées dans l'article (interview d'experts, phase d'évaluation, description du sport...).

Combiner mouvements analysés et vidéo du match réel Les sportifs ont souvent besoin d'analyser leur performance mais les critères statistiques la caractérisant sont beaucoup plus parlant lorsqu'ils sont associés à l'action à laquelle ils font référence. C'est pourquoi certains chercheurs s'attachent à intégrer les critères statistiques intéressants au sein des vidéos de jeu [15]. Toutefois il n'est parfois pas possible de représenter ces critères visuellement mais mettre la vidéo de l'action correspondante en parallèle peut s'avérer déjà concluant comme le fait l'interface BKViz [16], qui est une interface d'analyse de données au basket qui regroupe un grand nombre d'outils d'analyses différents avec en parallèle la vidéo de la phase de jeu correspondante. L'apport de cet outil réside dans le fait qu'il permet de naviguer entre les différentes données disponibles et les différents outils d'analyse d'une manière simple et efficace.

1.3 Source de données disponibles

Nous disposons actuellement des données pour 3 sports différents : le Football, le Basket et le Rugby. Pour chacun de ces sports ces données ont des formats différents dont chacun d'eux présente des avantages et des inconvénients :

- basket : 632 matchs de NBA. Chaque match est découpé en évènements. Chaque évènement contient la position des joueurs des deux équipes et du ballon.
- foot : données pour deux matchs pour les deux équipes. Nous avons également des données pour 3 autres matchs mais seulement avec une équipe.
- rugby : tous les matchs et entraînements du LOU depuis 2016.

Dans notre cas, étant donné la quantité et la qualité de données disponibles pour le basketball, nous nous concentrons sur celles-ci dans un premier temps.

Le jeu de données que nous utilisons est issu des données de l'entreprise Stats et de la technologie SportsVU. Celles-ci sont celles de 600 matchs de basket masculin en NBA. Entre les saisons 2013-2014 et 2016-2017, c'était SportsVU qui travaillait avec la NBA. L'année suivante, le championnat américain a fait le choix d'une autre technologie de video-tracking. Dans notre cas, ces données sont celles issues des matchs en play-by-play, c'est à dire action par action. Chacune est mise sous la forme d'un dictionnaire. Elles sont constituées des clés suivantes : visitor, gamedate, events, gameid et home [17] [18]. Le point qui nous intéresse est principalement la clé events qui se trouve elle aussi être un dictionnaire.

Sur la figure suivante est présentée la structure générale des données :

1.4 Objectifs

Mise au point et étude d'indicateurs de comportements collectifs. Le premier objectif consiste à rechercher des critères quantitatifs, définis à partir de l'analyse de données brutes, qui pourraient caractériser le comportement collectif d'une équipe de Basket. On cherchera ainsi à mettre en évidence des corrélations entre des modélisations mathématiques (comme les espaces de Voronoï) et des aspects qualitatifs (réalisation d'une passe par exemple) d'une phase de jeu. Il conviendra de réaliser un mini rapport

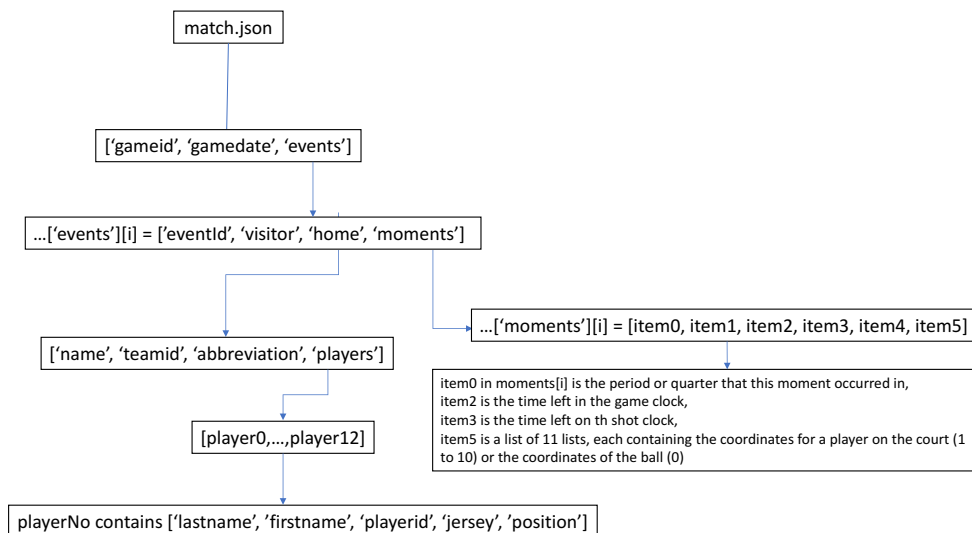


FIGURE 1 – Architecture des données de matchs de Basketball NBA

pour chaque critère testé qui contiendra la raison pour laquelle on a choisi ce critère, la démarche de test et une analyse des résultats. Ces éléments devront être versés sur un espace partagé.

Cet objectif sera atteint si les résultats trouvés (positifs ou négatifs) pourront intéresser.

Mise en forme de ces données pour les visualiser. Ce second objectif s'inscrit dans l'optique de rendre accessibles et compréhensibles les données récoltées par les capteurs. A ce jour, le PAR de l'année précédente a été confronté à des difficultés pour justifier son travail. La création d'un outil de visualisation améliorerait grandement ce point. On distingue deux phases dans cet objectif :

1. Proposer un type de représentation du traitement des données (sous *Python*),
2. Réaliser une application de visualisation complète exploitable par des experts sportifs non scientifiques (en *Javascript* a priori).

Dans notre cas, notre travail se concentre principalement sur la première phase. Pour des raisons de lacunes techniques et faute de temps, la seconde phase sera commanditée à un contact extérieur.

La représentation du traitement des données est une étape importante du projet. En effet, c'est l'interprétation de la représentation qui permettra de décider de l'intérêt ou non d'un critère. Puis la création d'une interface concrétisera ce travail en proposant un outil dédié aux spécialistes.

2 Résultats intermédiaires

2.1 Introduction

L'étude se concentrait sur le basketball étant donné la quantité et la qualité du jeu de données dont nous disposons (cf 1.3). A l'issue du RVP1 et de réunions ultérieures, il a été décidé de se concentrer principalement sur l'exploration de critères intéressants. De fait, la conception d'une application d'interface a été laissée de côté pour un travail ultérieur.

2.2 Démarche générale

Nous avons décidé de concentrer notre travail autour de l'étude des passes. L'idée serait de pouvoir visualiser, à un instant donné, toutes les possibilités de passes que le joueur en possession de la balle a. De manière visuelle, on souhaiterait mettre en évidence des zones du terrain où une passe serait possible. Comme mentionné dans la section 1.2.2 ce travail autour des passes est étroitement lié avec l'étude de l'occupation d'espace par les joueurs. Le principe de notre démarche est le suivant :

1. on souhaite dans un premier temps déterminer les zones du terrain contrôlées par les attaquants. En effet, dans l'objectif de prédire les passes, on peut commencer par une remarque évidente : les zones de réception de passe correspondent à des valeurs élevées, c'est à dire à des zones très bleues. Ceci s'interprète facilement : les zones de réception sont les zones où l'équipe attaquante est très influente. C'est le principe de la conservation de la balle : il est rare de tenter une passe vers une zone contrôlée par un adversaire.
2. Dans un second temps on cherchera à déterminer des couloirs de passe : trajectoire que la balle peut pourrait prendre sans être interceptée.
3. Dans un troisième temps on combinera les deux études afin de déterminer des zones de passes.

2.3 Démarche détaillée

On s'intéresse à l'occupation du terrain par les défenseurs et les attaquants. On ne prend pas en compte la position du ballon. On va chercher à découper le terrain en zones dans le but de déterminer des zones où les attaquants pourraient faire une passe en ayant une bonne chance de garder la balle sans prendre en compte les éventuelles interceptions.

2.3.1 Diagrammes de Voronoï

La première idée pour découper le terrain consiste à utiliser les diagrammes de Voronoï. Ces diagrammes découpent un plan en cellules à partir d'un ensemble de point appelés germes. Chaque cellule enferme un seul germe, et forme l'ensemble des points du plan plus proches de ce germe que de tous les autres.

Dans notre cas les germes sont les joueurs. On peut donc découper le terrain en associant à chaque joueur une zone lui appartenant. La figure 1 illustre ce découpage.

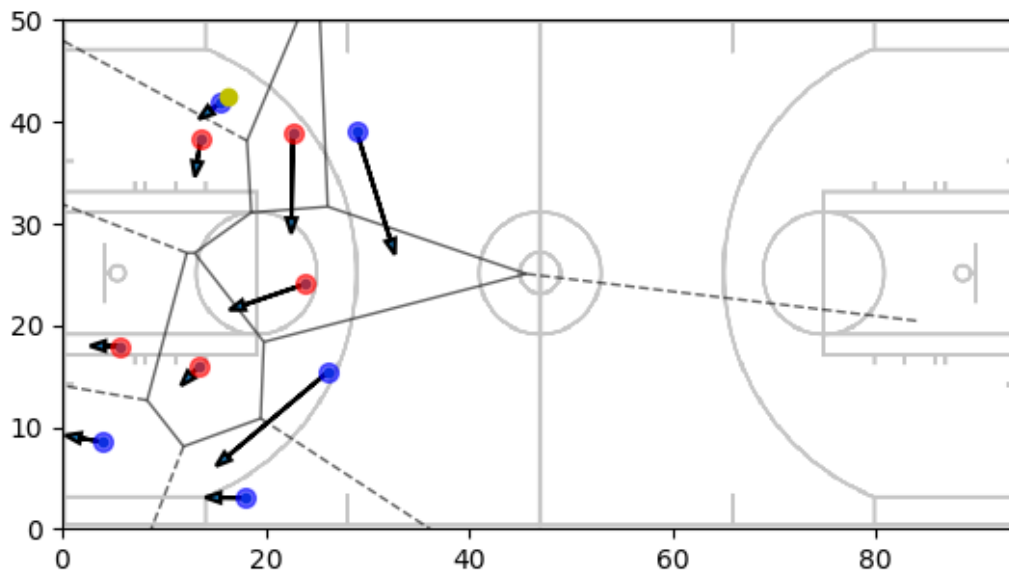


FIGURE 2 – Diagramme de Voronoï de joueurs de Basket. Les flèches noires représentent les vecteurs vitesses des joueurs (échelle agrandie).

Marc a réalisé différentes études à propos des espaces de Voronoï (voir détail section 1.2.1). Il avait remarqué que juste avant qu'une équipe marque au basket la surface moyenne des espaces de Voronoï de l'équipe qui attaque diminue au moment du tir. On essaiera dans notre cas de voir s'il y a des liens entre la variation des espaces de Voronoï et le moment où un joueur reçoit une passe.

2.3.2 Évaluation de l'occupation d'espace

Les diagrammes de Voronoï permettent une première approche simple pour le découpage de l'espace. Cependant au sein d'une zone associée à un joueur celui-ci contrôle plus les points proches de lui que les points loin. Ainsi, nous avons cherché des pistes pour *pondérer* l'influence des joueurs dans la détermination des espaces de Voronoï. Notre démarche est la suivante :

1. on réalise un maillage du terrain de basket (en réalité seulement la moitié du terrain car pour une phase de jeu au basket seule une moitié du terrain est concernée)
2. pour chaque maille on calcule le temps minimum que met un joueur de l'équipe attaquante à parvenir à cette maille, noté t_{att}
3. on calcule le temps minimum que met un joueur de l'équipe défendant à parvenir à ce point, noté t_{def}
4. on affecte la valeur $\delta = t_{att} - t_{def}$ à ce point,
5. on affiche une *heat-map* de l'ensemble des résultats.

Ainsi le terrain est découpé en valeur d'occupation de la manière suivante : plus un attaquant peut accéder vite à une maille par rapport aux défenseurs, plus il contrôle celle-ci, plus la quantité δ est grande.

Dans cette démarche, la méthode de calcul du temps minimum pour accéder à un point est très importante. Dans un premier temps, nous avons utilisé une méthode simple :

tous les joueurs peuvent rejoindre une maille avec une même vitesse peu importe leur vitesse initiale et leur position. La figure 3 illustre un résultat visuel obtenu grâce à cette démarche.

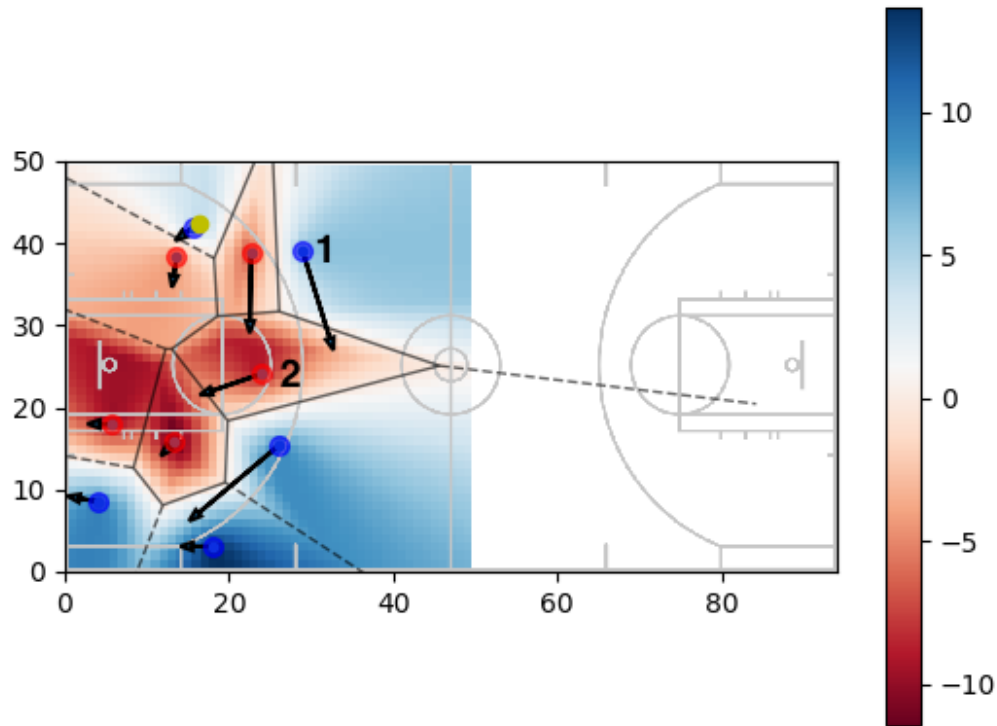


FIGURE 3 – Diagramme de Voronoï des joueurs avec valuation de l'espace : plus la couleur bleue est intense plus l'attaquant contrôle la zone, plus la couleur rouge est intense plus le défenseur contrôle la zone. Les flèches noires représentent les vecteurs vitesses des joueurs (échelle agrandie).

2.3.3 Prise en compte de l'inertie

La démarche permet de mieux se rendre compte de l'occupation de l'espace par les joueurs. Cependant elle ne prend pas en compte le fait que les joueurs sont en mouvement, or de manière intuitive, si un joueur court vite vers dans une direction il ne contrôle pas ou peu la zone derrière lui. On a donc cherché à prendre en compte la vitesse des joueurs dans le calcul du temps que mets un joueur pour rejoindre une passe. Suite à une discussion avec WOUTER BOS, nous avons conjointement développé un modèle reposant sur le principe de la gravitation. On considère pour chaque joueur J_i sa vitesse instantanée \vec{U}_{0i} , auquel on va appliquer une force constante en norme pour l'attirer vers la maille considéré. De fait, on trouvera une expression analytique pour t , le temps pour atteindre la maille. On obtient (voir détail en Annexe B) :

$$t^4 - \frac{4}{F^2} ((x_0 + u_0 t)^2 + ((y_0 + v_0 t))^2) = 0 \quad (1)$$

Ce modèle permet de prendre en compte la vitesse initiale. Il permet de trouver des trajectoires plus naturelles dans le mouvement qui rapproche un joueur d'un point. Afin de mieux comprendre la trajectoire qu'un joueur réalise pour rejoindre une maille grâce à cette méthode on illustre figure 4 la trajectoire d'un joueur (point bleu) pour rejoindre

un point (point rouge) en ayant un vecteur vitesse initiale (flèche verte) et subissant une force constante pendant la trajectoire (flèche noire).

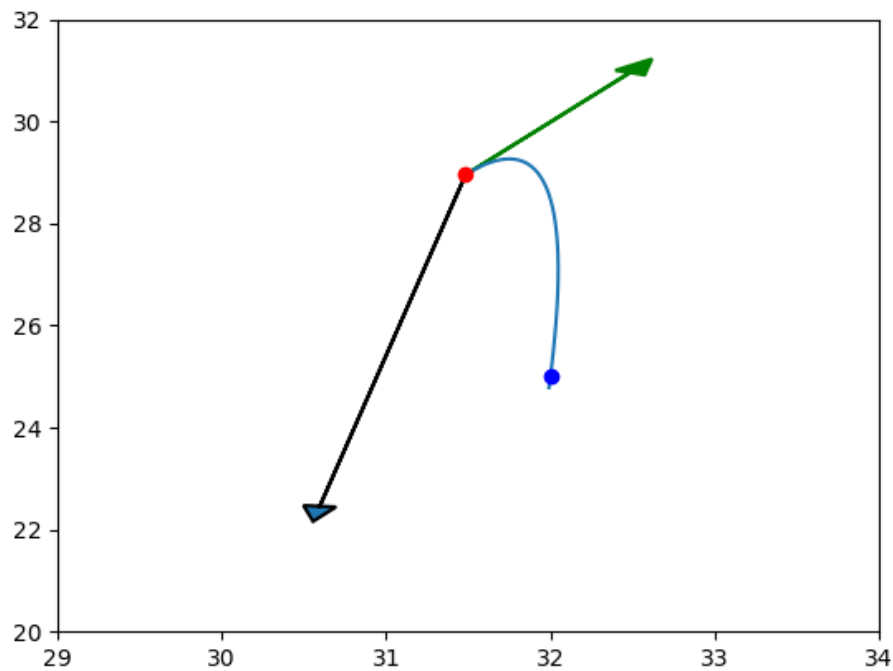


FIGURE 4 – Trajectoire d'un joueur (point bleu) à l'aide de la méthode décrite précédemment

L'*heat-map* résultant de cette méthode de calcul du temps est donnée figure 5 :

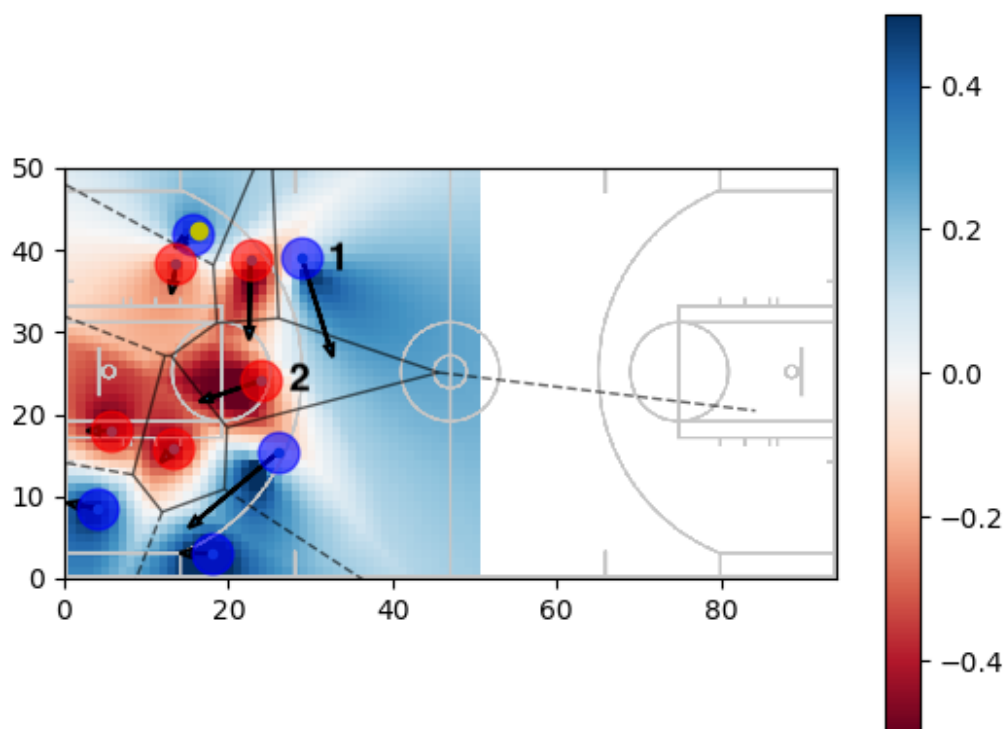


FIGURE 5 – *Heat-map* obtenue en prenant en compte la vitesse des joueurs. Les traits noirs représentent les diagrammes de Voronoï qui eux ne prennent pas en compte la vitesse des joueurs. Les flèches noires représentent les vecteurs vitesses des joueurs (échelle agrandie).

Afin de mieux visualiser ce critère nous avons calculé l'*heat-map* sur un ensemble de moments qui rassemblés constituent une phase de jeu. Puis nous avons assemblé ces images ce qui nous a permis d'obtenir une vidéo qui superpose le mouvement des joueurs et les zones occupées par les joueurs.

2.4 Calcul des aires des zones contrôlées

Donner une valeur d'occupation de l'espace par les joueurs, en prenant en compte leur inertie a permis de mettre en évidence de manière visuelle de quelle manière les joueurs contrôlent le terrain. L'observation de ce critère sur une phase de jeu grâce à la vidéo, nous a invité à comparer les surfaces des zones contrôlées par les deux équipes et à regarder s'il n'existe pas un lien entre la variation de ces surfaces et le moment où une équipe fait une passe.

Dans un premier temps nous avons simplement utilisé un compteur qui, lors du parcours des différentes mailles du terrain, est incrémenté si cette maille est occupée par les attaquants. Notons que dans cette première idée on ne prend pas en compte l'évaluation $\delta = t_{att} - t_{def}$, on cherche à voir si une maille est occupée par un attaquant ou un défenseur en regardant si δ est positif ou non. Notons également que le calcul de δ est effectué à l'aide de la démarche décrite dans la section 2.3.3 et prend donc en compte l'inertie des joueur.

Dans un second temps nous avons également cherché à voir si de manière individuelle, il existe un lien entre le fait qu'un attaquant reçoit une passe et l'évolution de sa surface par

rapport à celle de ces coéquipiers. Le principe de calcul est le même que précédemment : on parcourt les différentes mailles et si elle appartient à un des attaquants on incrémente son compteur de un.

2.5 Analyse des résultats

Dans cette partie, nous analysons les différents résultats obtenus grâce à notre démarche expliquée précédemment.

2.5.1 Influence de l'inertie sur les espaces de Voronoï

Les figures 6 et 7 permettent de comparer l'évaluation de l'occupation de l'espace par les joueurs en prenant en compte ou non l'inertie des joueurs. Notons que sur la figure sont représentés :

- les joueurs et la balle,
- les espaces de Voronoï (traits noirs),
- les vecteurs vitesses des joueurs agrandis pour une meilleure visualisation,
- la *heat-map* correspondant à la valeur de contrôle de l'espace par les joueurs (plus une zone est bleue, plus un attaquant la contrôle, plus une zone est rouge plus une zone la contrôle).

On remarque que sur la figure 6 les zones bleues et rouges sont délimitées par les espaces de Voronoï ce qui n'est pas le cas sur la figure 7 ce qui montre que prendre en compte la vitesse des joueurs apporte un regard différent sur l'occupation d'espace par les joueurs. Si on regarde de plus près cette approche paraît plus cohérente. En effet si on se concentre sur les joueurs 1 et 2 et sur la zone délimitée par le cercle vert, on constate que le joueur 1 se dirige vers cette zone avec une vitesse élevée tandis que le joueur 2 se dirige à l'opposé de cette zone avec une vitesse modérée. Ainsi de manière intuitive ce cercle vert est plus contrôlé par le joueur 1 que par le joueur 2, donc par un attaquant ce qui veut dire la zone devrait être bleue. Or sur la figure 6 la zone est colorée en rouge alors que sur la figure 7 la zone est colorée en bleu. La prise en compte de l'inertie pour étudier l'occupation d'espace est donc judicieuse et permet de mieux rendre compte de l'occupation réelle des joueurs.

Notons que les zones blanches ou très claires sont des zones où l'attaque et la défense ont autant d'influence.

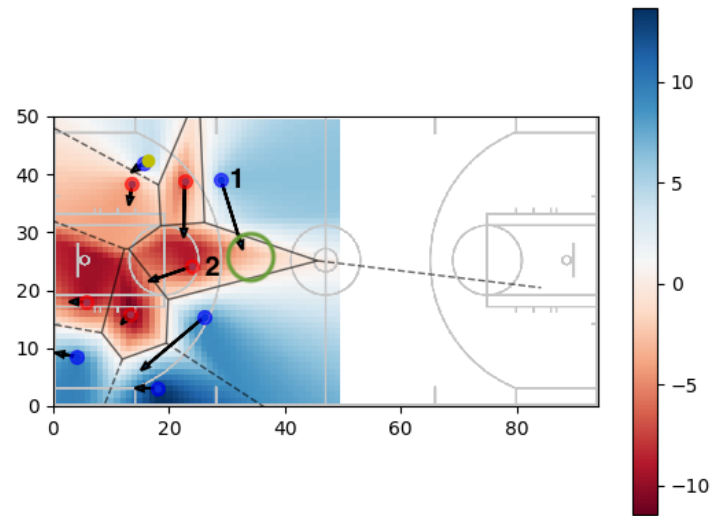


FIGURE 6 – *Heat-map* qui ne prend pas en compte l’inertie des joueurs

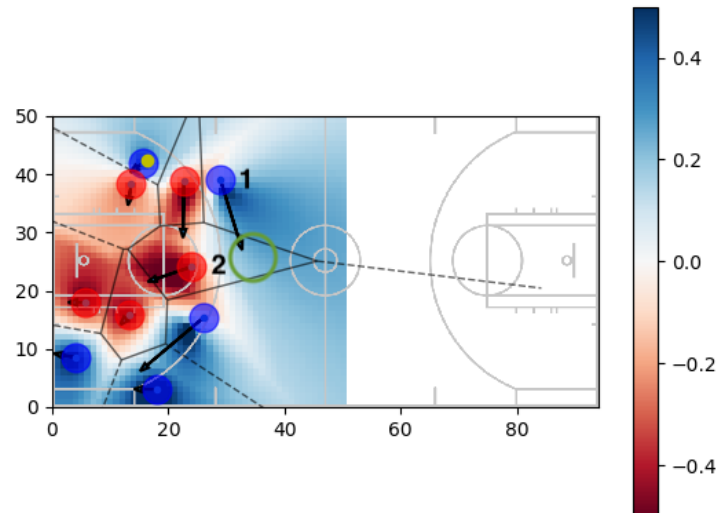


FIGURE 7 – *Heat-map* calculée en prenant en compte l’inertie des joueurs

On remarque que sur la figure 6 les zones bleues et rouges sont délimitées par les espaces de Voronoï ce qui n’est pas le cas sur la figure 7 ce qui montre que prendre en compte la vitesse des joueurs apporte un regard différent sur l’occupation d’espace par les joueurs. Si on regarde de plus près cette approche paraît plus cohérente. En effet si on se concentre sur les joueurs 1 et 2 et sur la zone délimitée par le cercle vert, on constate que le joueur 1 se dirige vers cette zone avec une vitesse élevée tandis que le joueur 2 se dirige à l’opposé de cette zone avec une vitesse modérée. Ainsi de manière intuitive ce cercle vert est plus contrôlé par le joueur 1 que par le joueur 2, donc par un attaquant ce qui veut dire la zone devrait être bleue. Or sur la figure 6 la zone est colorée en rouge alors que sur la figure 7 la zone est colorée en bleu. La prise en compte de l’inertie pour étudier l’occupation d’espace est donc judicieuse et permet de mieux rendre compte de l’occupation réelle des joueurs.

Notons que les zones blanches ou très claires sont des zones où l’attaque et la défense ont autant d’influence.

2.5.2 Le lien avec la prédiction de passe

Interprétation visuelle Grâce à la vidéo réalisée sur une phase de jeu on a pu observé des passes entre joueurs tout en visualisant leur occupation du terrain. Les figures 8 et 9 représentent une situation de passe du joueur 1 vers le joueur 2. La figure 8 représente la situation juste avant la passe et la figure 9 la situation juste après la passe.

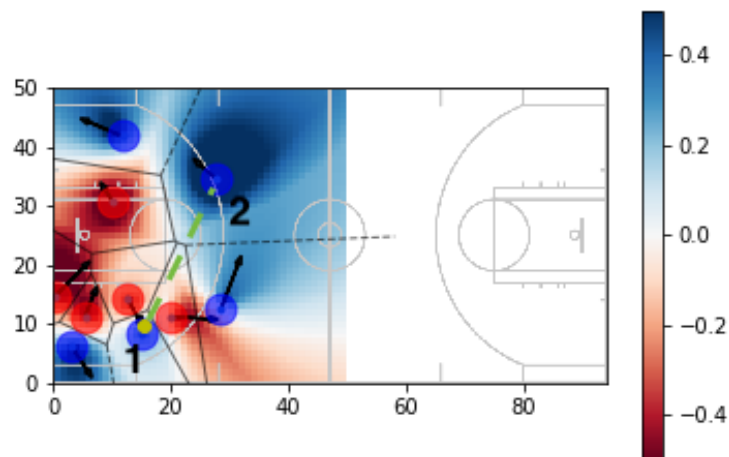


FIGURE 8 – Image du terrain avant que le joueur 1 fasse la passe au joueur 2 (trajectoire en pointillés verts)

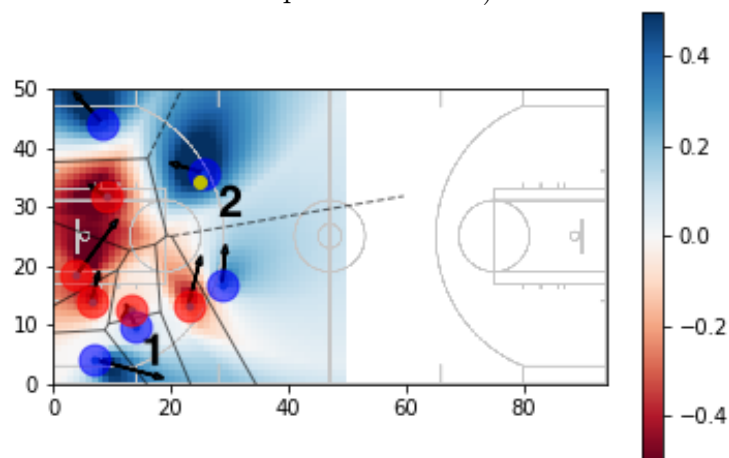


FIGURE 9 – Image du terrain au moment où le joueur 2 reçoit la balle

Lors de cette passe on constate deux choses. La première est que la passe a été effectuée d'une zone peu contrôlée par les attaquants vers une zone avec une forte valeur de contrôle. Ceci peut paraître logique puisque que le joueur en possession du ballon reçoit nécessairement plus de pression de la part des défenseurs que ses coéquipiers et il va donc chercher à faire la passe à un joueur en meilleure position que lui. Cette situation révèle une passe pour laquelle le receveur contrôle une très grande zone autour de lui mais dans la vidéo ceci n'est pas toujours le cas.

La deuxième chose qu'on observe est que la trajectoire de la passe traverse que des zones blanches ou claires. Ce ne sont donc pas des zones sécurisées mais pas non plus des zones contrôlées par les adversaires.

Ces deux remarques permettent de soulever le fait qu'une passe peut-être plus ou moins risquée (risque d'interception) et avec une plus-value plus ou moins grande pour

l'équipe (receveur libre et en bonne position sur le terrain).

Surfaces occupées par les joueurs et passes L'interprétation visuelle précédente a permis de soulever l'idée qu'un joueur fait intuitivement une passe d'un endroit peut contrôlé vers un endroit mieux contrôlé. À l'aide du calcul des surfaces contrôlées par les équipes et les joueurs (voir section 2.4) nous allons essayé de voir s'il existe un lien entre occupation d'espace et les situations de passes.

Premièrement, on regarde l'évolution de l'aire occupée par les attaquants et les défenseurs au cours de la phase de jeu. Les traits verticaux noirs situent une passe et le trait vertical vert situe le shoot final de l'action. La figure 10 présente cette évolution. Cette évolution ne permet pas clairement de mettre clairement en évidence un lien entre passe et occupation de l'espace. On peut néanmoins faire une remarque : les 3 passes sont précédées d'un gain en surface contrôlée, on pourrait ainsi penser qu'avant de faire une passe il faut que les joueurs se démarquent et libèrent donc de l'espace. Il pourrait être intéressant d'essayer de vérifier ce constat sur un grand nombre de phase de jeu.

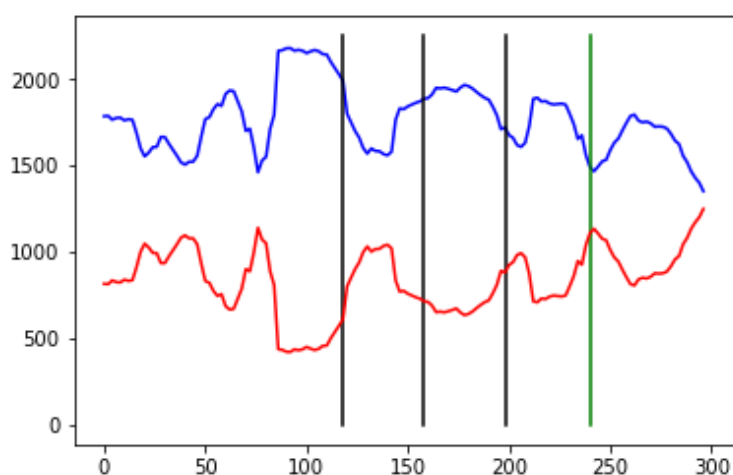


FIGURE 10 – Évolution de l'aire occupée par les attaquants (en bleu) et par les défenseurs (en rouge) en fonction du temps

Deuxièmement, on cherche à regarder l'évolution de l'aire contrôlée par les attaquants individuellement pour voir le comportement avant de recevoir la balle et pour voir si une passe se fait d'une zone peu contrôlée vers une zone mieux contrôlée comme nous l'avons suggéré précédemment.

La figure 11 représente l'évolution de l'aire occupée par chaque attaquant en fonction du temps.

On remarque que pour chaque passe, le passeur contrôle une plus petite surface que le receveur. La différence de contrôle peut néanmoins être faible mais on réalise quand même cette observation. Il semblerait ainsi que la remarque faite dans la section précédente se vérifie, il faudrait effectuer des tests sur un grand nombre de phase de jeu pour réellement valider cette remarque. On note également que la passe qui précède le tir est une passe vers un joueur qui possède une très grande surface de contrôle, ainsi une telle passe possède une bonne plus-value.

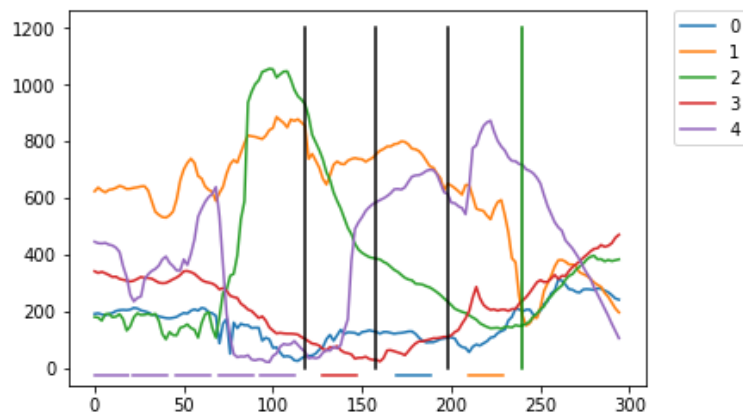


FIGURE 11 – Évolution de l'aire occupée par chaque attaquant en fonction du temps.
Les traits au niveau du zéro indiquent quel joueur possède la balle

Bilan L'approche sur l'occupation d'espace permet de révéler des zones potentielles où les joueurs peuvent recevoir la balle. Néanmoins Cependant, assurer que la balle se dirige vers une zone contrôlée par un joueur attaquant est important mais pas suffisant. Il faut s'assurer que la balle ne soit pas interceptée dans son parcours vers le joueur visé. Notre représentation ne permet pas de donner qualitativement de règle sur ce critère. C'est une des limites de cette visualisation et de notre modèle en 2 dimensions. C'est donc un point essentiel à travailler pour prédire une passe.

De plus nos observations afin de lier passes et occupation d'espace n'ont été faites que sur de faibles quantités d'événements. Afin de voir s'il existe un lien il faut réaliser des études statistiques sur un grand nombre de phases.

2.5.3 Remarques diverses

Plusieurs remarques sont possibles par rapport à notre travail :

1. Dans notre évaluation de l'occupation de l'espace nous prenons seulement en compte le temps d'un attaquant (celui qui arrive le plus vite) et le temps d'un défenseur (celui qui arrive le plus vite) pour le calcul de δ or ceci ne prend pas en compte le fait que même si un attaquant arrive plus vite que les défenseurs dans une zone il ne la contrôle pas de la même manière s'il y a un ou deux défenseurs proches de cette zone.
2. Une démarche possible (proche de celle utilisée dans [6] autre que la notre serait d'à partir d'un instant t , calculer les zones possibles où les joueurs et le ballon peuvent se trouver 1 seconde plus tard et de voir les d'intersections entre la zone du ballon et des attaquants pour déterminer des zones de passes, entre la zone du ballon et les défenseurs pour les zones possibles d'interceptions.
3. Il serait intéressant de donner une valeur aux différentes zones du terrain notamment par rapport à la position par rapport au panier. En effet dans notre approche on regarde seulement si un joueur est libre ou non, or un joueur pourrait se trouver au milieu du terrain tout en étant libre mais une passe vers ce joueur ne serait pas judicieuse pour autant.

2.6 Espace partagé

Dans le cadre de notre travail, nous avons cité comme tâche la *Gestion du code créé*. Ce travail correspond à la mise en commun des codes créés, en Python ou en Javascript, et à permettre une maintenance et une évolution simplifiée.

Ainsi, un espace partagé a été créé sur la plateforme GITHUB. Ce dernier est disponible à l'adresse suivante : <https://github.com/AmigoCap/MecaFootCo>. Il présente les éléments suivants :

- une présentation du projet et un récapitulatif des travaux réalisés,
- une bibliographie commentée,
- l'ensemble des codes édités et une explication associée à chacun.

Une capture d'écran de l'interface de visualisation est présentée sur la figure 12.

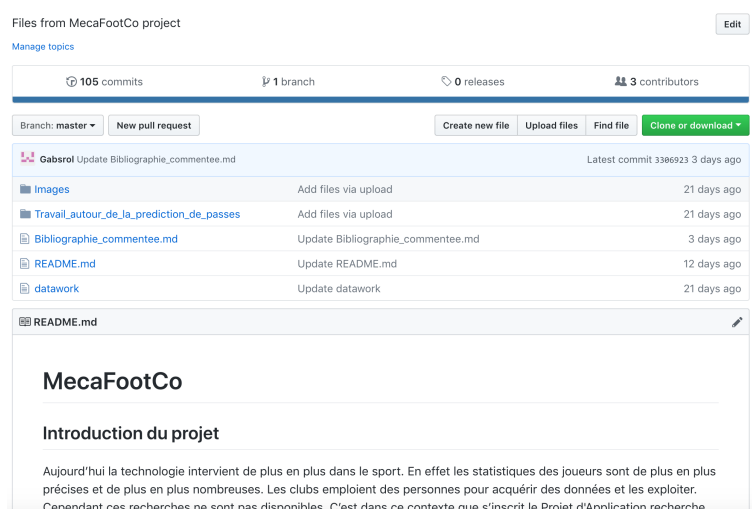


FIGURE 12 – Capture d'écran de l'espace partagé

2.7 Travaux à venir

Approfondir calculs des aires L'étude de l'évolution des aires en lien avec les passes n'a été faite que sur une phase de jeu. Comme mentionné dans la section ?? afin de confirmer ou non nos différentes observations sur une phase de jeu on va essayer de réaliser des tests sur un grand nombre de phases de jeu.

Étudier la valeur d'occupation de la zone où une passe est faite Il sera intéressant de regarder pour chaque passe la valeur δ (voir section 2.3.3) au niveau de la zone de réception afin de voir les valeurs de l'occupation d'espace par le receveur.

Déterminer les couloirs de passes possibles Comme énoncé dans la partie *Le lien avec la prédiction de passe* (2.5.2), notre représentation n'est pas en mesure de nous informer suffisamment sur les trajectoires possibles pour une passe. Or, c'est un critère essentiel pour les prédire. Ainsi, nous devons trouver un modèle permettant de caractériser ce critère. Nous avons pour idée de proposer un modèle simple dont la démarche pourrait être la suivante :

1. on fixe la norme de la vitesse de la balle u à tout instant,

2. pour chaque point de notre maille, on détermine si la balle arrive avant un joueur adverse ou non (valeurs 0 ou 1),
3. on colorie en noir si 0, en blanc si 1.

Ainsi on obtiendrait une représentation des couloirs de passe possibles pour une balle évoluant à la vitesse u . Cependant, ce modèle est très simple car on considère la vitesse de la balle constante. Il sera à améliorer au fur et à mesure des observations.

En combinant ce modèle avec le modèle de zone d'influence on pourra déterminer des zones de passe comme étant l'intersection d'un couloir de passe et une zone contrôlée par un coéquipier (zone bleue sur les images). Il conviendra donc de définir à partir de quelle intensité de bleu on considère que la zone est contrôlée par un joueur.

3 Conclusion

Ce livrable intermédiaire a permis d'exposer les travaux déjà réalisés au sein de ce domaine de recherche. Beaucoup d'éléments ont déjà été étudiés (occupation de l'espace, possibilité de passe et analyse de tir, lien entre sport et sciences). Toutefois, il reste des voies à explorer, notamment pour le basket, étant donné la quantité et qualité des données disponibles.

Dans cette optique, nous nous sommes concentrés sur la prédiction de passe pour ce sport, qui relève d'un enjeu important pour chaque phase de jeu. Ainsi, nous tentons de créer des métriques, notamment pour l'occupation de l'espace, possiblement utiles à la prédiction de passe. A ce stade, l'occupation de l'espace, influencée par l'inertie des joueurs, a pu être caractérisée par une méthode incluant un modèle gravitationnel. Une visualisation par le biais d'une *heat-map* a été proposée. Elle permet la comparaison avec un outil mathématique connu : les espaces Voronoï. Des premières interprétation et analyses ont pu être réalisées suites à la simulation de ces *heat-map* sur des actions. On obtient un résultat intéressant pour l'étude des zones d'influence, mais insuffisant pour prédire correctement une passe. En effet, les zones d'influences ne décrivent pas ce qui adviendra de la balle lors de son parcours, c'est un point essentiel à travailler. Conjointement, nous analyserons plus en détail les indicateurs à l'endroit où une passe est réalisée.

Annexe A

FICHE DE LANCEMENT DE PROJET



La fiche de lancement du projet est un **document de synthèse** qui a pour but de présenter le projet de manière claire et **compréhensible par tous**. Elle définit le **contour du projet**, les **objectifs visés**, les **cibles** et les **contraintes** s'exerçant sur le projet. Elle nécessite de prendre du recul, d'avoir une vision globale et synthétique du projet et de son environnement, ainsi que des éléments clés du projet.

La rédaction de cette fiche est une étape primordiale qui conditionne le bon lancement du projet. Elle est commune aux trois types de projets proposés à l'École Centrale de Lyon : PE, PAi et PAr.

Fiche d'identité	Titre et éventuel acronyme : MecaFoot Co
	Tuteur ou Commanditaire : Wouter BOS
Contexte	Origine du besoin : Aujourd'hui la technologie intervient de plus en plus dans le sport. En effet les statistiques des joueurs sont de plus en plus précises et de plus en plus nombreuses. Les clubs emploient des personnes pour acquérir des données et les exploiter. Cependant ces recherches ne sont pas disponibles. C'est dans ce contexte que s'inscrit le PAr MécaSportCo dont le but est d'analyser des données sportives pour trouver un ou plusieurs critères qui pourrait être utiles aux experts.
	Enjeux : <ul style="list-style-type: none"> - Améliorer la performance collective d'une équipe - Effectuer un travail de recherche dans le cadre de Sciences 2024 - Expliquer par des facteurs quantitatifs des événements de jeu.
Objectifs	Objectif général : Progresser dans l'exploration des différentes caractérisations possibles des comportements des sportifs à partir de l'exploitation de différentes données ainsi que dans le développement d'une interface visuelle permettant de communiquer ces données.
	Indicateurs mesurables de réussite : Découverte d'indicateurs intéressants ou de méthodes de prédiction (prédiction de la passe par exemple).
	Nature du livrable principal : Rapport ou article (à définir)
Périmètre	Acteurs : Wouter Bos (tuteur & commanditaire), Romain Vuillemot (tuteur & commanditaire), Marc Mattis (ancien élève sur le PA), Gabin Rolland et Nathan Rivière (élèves actuellement sur le PA).
	Ressource : Données sportives collectées précédemment.
	Environnement et interfaces du projet : Environnement informatique (Python,...).
Contraintes	Coûts : Coûts de 2 ingénieurs pour un volume horaire de 270€ chacun.
	Délais Livrable à rendre à la fin de l'année scolaire de 2A.
	Autres contraintes :

Annexe B

Determine the closest player to a point

Wouter Bos

December 2018

We want to compute, for a moving player on a domain how much time it takes to reach the point $(0, 0)$.

At the initial time the velocity of the player at position $\mathbf{x}(t=0) = (x_0, y_0)$ is $\mathbf{u}(x_0, y_0, t=0) = (u_0, v_0)$. We will consider that the player will use a constant force (per unit mass) in a given direction, of strength $|\mathbf{F}|^2 = F_x^2 + F_y^2$. This is an assumption which allows to find a simple analytical solution. In particular, it allows to consider the two directions separately.

Newton's law writes

$$d_t^2 x = F_x \quad (1)$$

so that we have

$$x(t) = x_0 + u_0 t + \frac{1}{2} F_x t^2. \quad (2)$$

We evaluate this expression at $x = 0$, and want to determine at which time this point is reached. Let us first determine the force per mass F_x ,

$$F_x = -2 \frac{x_0 + u_0 t}{t^2}. \quad (3)$$

and similarly

$$F_y = -2 \frac{y_0 + v_0 t}{t^2}. \quad (4)$$

Since $|\mathbf{F}|^2 = F_x^2 + F_y^2$, we have,

$$F^2 = \frac{4}{t^4} ((x_0 + u_0 t)^2 + (y_0 + v_0 t)^2). \quad (5)$$

Yielding the fourth order polynomial for t ,

$$t^4 - \frac{4}{F^2} ((x_0 + u_0 t)^2 + (y_0 + v_0 t)^2) = 0. \quad (6)$$

This equation has formally 4 solutions. However, only one of these is the physical time for a player to reach the origin. The constraints to choose the correct solution are that the time needs to be positive and real.

In practice, in a script, one can evaluate all four analytical solutions of this equation and choose the smallest positive real solution. The adjustable parameter is the value of F . It should be of order of magnitude $1 < F < 10 \text{ ms}^{-2}$.

Références

- [1] Benjamin Kadoch, Wouter J. T. Bos, and Kai Schneider. Directional change of fluid particles in two-dimensional turbulence and of football players. *Physical Review Fluids*, 2(6) :064604, June 2017.
- [2] Dan Cervone, Luke Bornn, and Kirk Goldsberry. NBA Court Realty. page 8, 2016.
- [3] Thorsten Breitzkreutz Manuel Stein, Halldór Janetzko and al. (PDF) Director's Cut : Analysis and Annotation of Soccer Matches.
- [4] Javier Fernandez and Luke Bornn. Wide Open Spaces : A statistical technique for measuring space creation in professional soccer. page 19, 2018.
- [5] Gennady Andrienko, Natalia Andrienko, Guido Budziak, Jason Dykes, Georg Fuchs, Tatiana von Landesberger, and Hendrik Weber. Visual analysis of pressure in football. *Data Mining and Knowledge Discovery*, 31(6) :1793–1839, November 2017.
- [6] Chan-hyun Kang, Jung-rae Hwang, and Ki-joune Li. Trajectory Analysis for Soccer Players. In *Sixth IEEE International Conference on Data Mining - Workshops (ICDMW'06)*, pages 377–381, Hong Kong, China, 2006. IEEE.
- [7] Alejandro Benito Santos, Roberto Theron, Antonio Losada, Jaime E. Sampaio, and Carlos Lago-Peñas. Data-Driven Visual Performance Analysis in Soccer : An Exploratory Prototype. *Frontiers in Psychology*, 9, 2018.
- [8] Yisong Yue, Patrick Lucey, Peter Carr, Alina Bialkowski, and Iain Matthews. Learning Fine-Grained Spatial Models for Dynamic Sports Play Prediction. In *2014 IEEE International Conference on Data Mining*, pages 670–679, Shenzhen, China, December 2014. IEEE.
- [9] Manuel Stein, Thorsten Breitzkreutz, Johannes Haussler, Daniel Seebacher, Christoph Niederberger, Tobias Schreck, Michael Grossniklaus, Daniel Keim, and Halldor Janetzko. Revealing the Invisible : Visual Analytics and Explanatory Storytelling for Advanced Team Sport Analysis. In *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, pages 1–9, Konstanz, October 2018. IEEE.
- [10] Kirk Goldsberry. CourtVision : New Visual and Spatial Analytics for the NBA. page 7, 2012.
- [11] Dan Cervone, Alexander D'Amour, Luke Bornn, and Kirk Goldsberry. Predicting Points and Valuing Decisions in Real Time with NBA Optical Tracking Data. page 9, 2014.
- [12] Amandine Aftalion, Louis-Henri Despaigne, Alexis Frentz, Pierre Gabet, Antoine Lajouanie, Marc-Antoine Lorthiois, Lucien Roquette, and Camille Vernet. How to identify the physiological parameters and run the optimal race. *MathematicS In Action*, 7(1) :1–10, 2016.
- [13] Charles Perin, Romain Vuillemot, C D. Stolper, J T. Stasko, J Wood, and Sheelagh Carpendale. State of the Art of Sports Data Visualization. *Computer Graphics Forum*, 37, June 2018.
- [14] Charles Perin, Romain Vuillemot, and Jean-Daniel Fekete. SoccerStories : A Kick-off for Visual Soccer Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12) :2506–2515, December 2013.

- [15] M. Stein, H. Janetzko, A. Lamprecht, T. Breitzkreutz, P. Zimmermann, B. Goldlücke, T. Schreck, G. Andrienko, M. Grossniklaus, and D. A. Keim. Bring It to the Pitch : Combining Video and Movement Data to Enhance Team Sport Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1) :13–22, January 2018.
- [16] Antonio G. Losada, Roberto Therón, and Alejandro Benito Santos. BKViz : A Basketball Visual Analysis Tool. *IEEE Computer Graphics and Applications*, 36 :58–68, November 2016.
- [17] Rajiv Shah. Exploring NBA SportVu Movement Data.
- [18] Savvas Tjortjoglou. How to Track NBA Player Movements in Python, August 2015.
- [19] Laure Cailloce. Quand les maths se mêlent de sport.
- [20] Bastien L. Sports Analytics : top des startups d’analyse de données sportives, June 2018.
- [21] Michael C. Rumpf. Performance analysis.