

« Automatic event detection in football using tracking data »

Dans cet article les auteurs expliquent que les données utilisées habituellement pour analyser les matchs de football sont souvent collectées à la main, ce qui fait qu'on les a seulement pour quelques tournois. L'article propose une nouvelle méthode pour collecter automatiquement ces données grâce à un algorithme spécifique. Cet algorithme fonctionne en deux étapes : la première sert à déterminer qui a le ballon à chaque instant du match et à analyser la position des joueurs quand le jeu est arrêté, et la seconde sert à identifier les différents événements du match, comme les changements de possession du ballon ou les fautes, en utilisant les règles du football. Cette méthode est très efficace, avec plus de 90% des événements bien identifiés comparés à une collecte manuelle, et elle pourrait aider à faire des analyses plus détaillées des matchs.

1) Sources des données

Les données proviennent de trois fournisseurs différents — Track160, Tracab, et Hawk-Eye — utilisées pendant des matchs de la Coupe du Monde des Clubs de la FIFA et de la Bundesliga. Ces données enregistrent les positions des joueurs et du ballon en 2D, avec une fréquence de 25 Hz. L'état du jeu est aussi suivi, indiquant si le ballon est en jeu ou pas.

Pour évaluer la précision des événements détectés par l'algorithme, les chercheurs ont utilisé les données officielles collectées par Sportec Solutions pour ces matchs. Ces données officielles comprennent des détails précis comme le temps et les joueurs impliqués dans chaque événement.

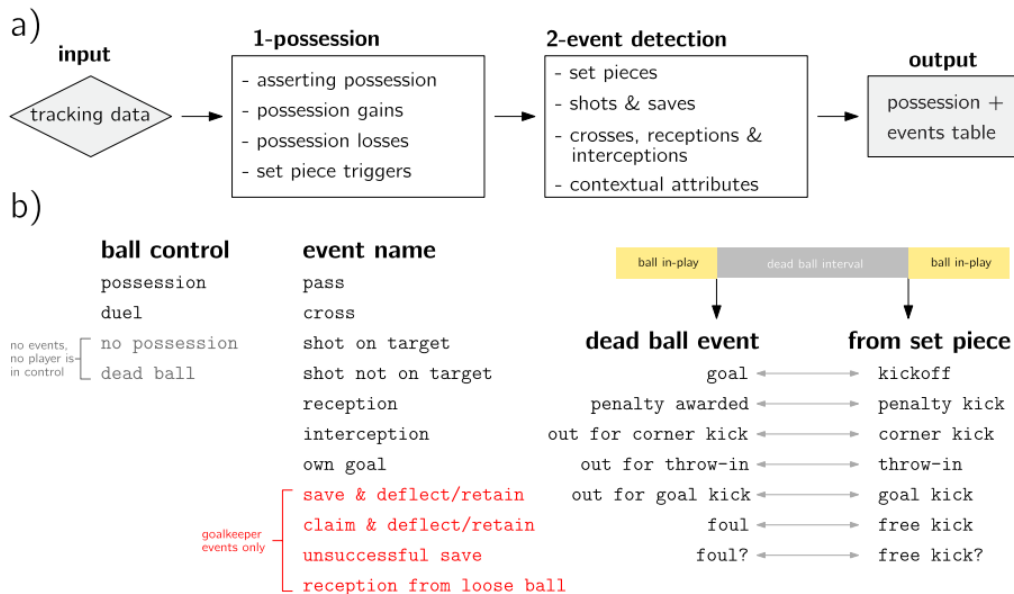
2) Comment ça marche ?

Le processus de détection automatique des événements commence par déterminer qui a la possession du ballon et la configuration des joueurs lors des interruptions de jeu. Ensuite, un arbre de décision basé sur les règles du football permet d'identifier les différents types d'événements de jeu.

L'algorithme produit un tableau pour chaque image du match, détaillant les événements qui s'y sont produits, qui inclut des informations sur le temps, les joueurs impliqués, et des attributs comme le contrôle du ballon, le type d'événement, et si l'événement découle d'une situation arrêtée ou non. Les différents contrôles de ballon peuvent être : ballon mort, sans possession, possession, ou duel. L'algorithme se concentre sur des événements spécifiques comme les passes, les tirs, les réceptions, et les interceptions, qui peuvent être identifiés sans apprentissage machine, simplement en appliquant des règles.

Des attributs spéciaux sont ajoutés pour les événements relatifs aux situations de balle morte, comme les corners, les sorties de but, les touches, les fautes, les penalties et les buts. Des détails supplémentaires pour reprendre le jeu après une interruption sont aussi identifiés, comme pour un corner ou un coup franc. Enfin, l'article mentionne que l'algorithme peut se

tromper parfois à cause des imprécisions dans les données de suivi, et une section supplémentaire dans une ressource en ligne donne plus de détails à ce sujet.



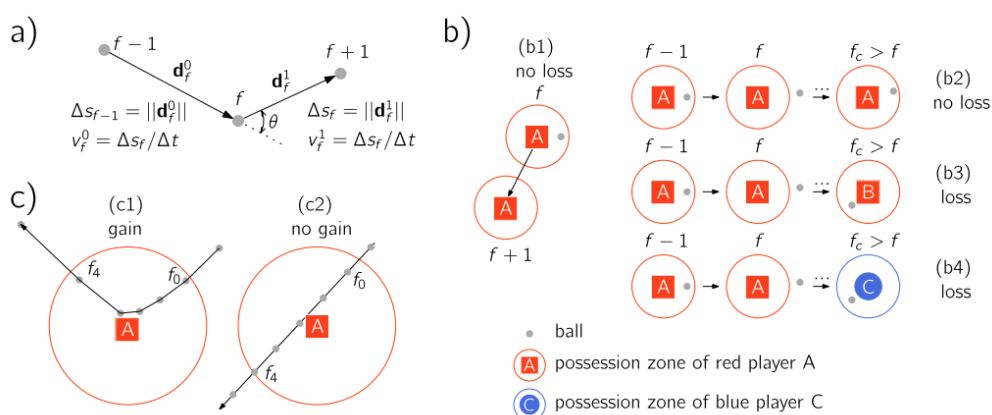
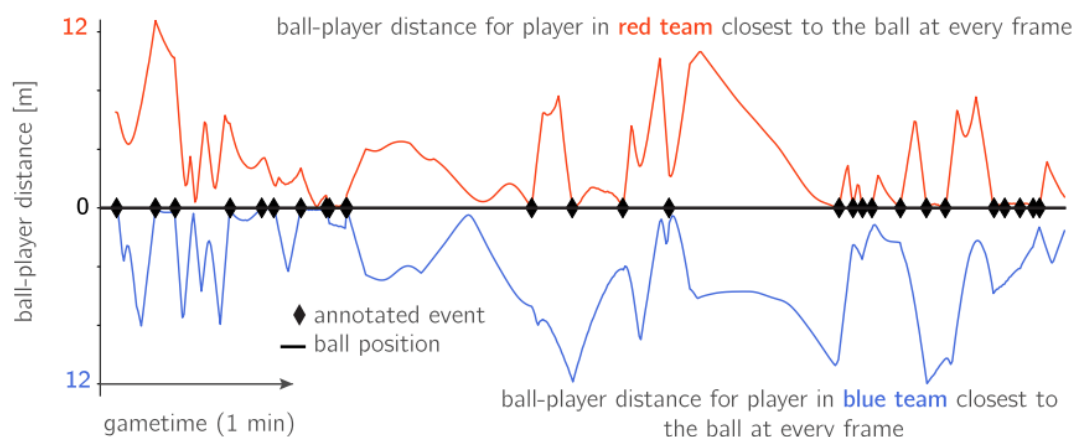
Pour la possession : Les auteurs expliquent comment l'algorithme détermine qui a la possession du ballon en utilisant les données de suivi lors d'un match de football. Le concept de "zone de possession" (PZ) est introduit : c'est une zone circulaire autour de chaque joueur. Si le ballon entre dans cette zone, le joueur est considéré comme ayant la possession. Il y a aussi une "zone de duel" (DZ) pour identifier quand deux adversaires sont en compétition pour le ballon.

L'algorithme vérifie chaque image pour voir si les conditions de possession ou de duel sont remplies, en donnant la priorité au duel si les deux conditions sont rencontrées. Pour chaque image où une possession ou un duel est identifié, l'algorithme enregistre plusieurs données comme la distance des joueurs au ballon, le déplacement et la vitesse du ballon.

Pour déterminer les changements de possession, l'algorithme regarde si le ballon quitte la zone de possession d'un joueur et si le mouvement du ballon dépasse un seuil spécifique. Il vérifie aussi si le joueur continue d'être impliqué dans les phases de jeu suivantes.

Concernant les gains de possession, l'algorithme cherche à confirmer le contact réel avec le ballon, en se basant sur les changements de direction et de vitesse du ballon. Si le ballon change de direction ou de vitesse durant une séquence donnée, cela indique un gain de possession. Si ni la direction ni la vitesse ne changent, les images correspondantes sont

considérées comme des faux positifs et sont exclues de l'analyse de possession.



Les déclencheurs de phases arrêtées ?

Cette section décrit comment l'algorithme intègre les informations de possession pour activer les déclencheurs des phases de jeu arrêtées, en fonction de la position des joueurs sur le terrain lorsque le jeu est interrompu. L'emplacement des joueurs permet à l'algorithme de déterminer quel type d'événement de jeu arrêté va reprendre le jeu. Déclencheur de balle morte : Le jeu est considéré comme arrêté si le ballon est statique, ce qui est signalé par un indicateur booléen ou l'absence de données de suivi du ballon.

Déclencheur de coup d'envoi : Les joueurs doivent être positionnés dans leurs moitiés de terrain respectives, avec au moins un joueur proche du centre du terrain, conforme à la Loi 8 de l'IFAB.

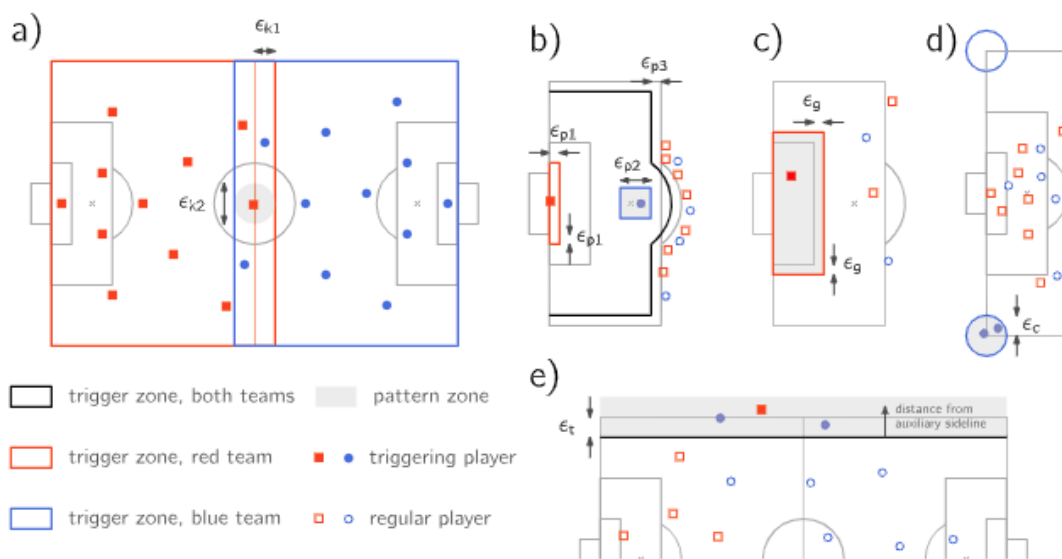
Déclencheur de penalty : Un joueur doit être sur sa ligne de but entre les poteaux et un adversaire doit se trouver dans une zone spécifique près du point de penalty, avec les autres joueurs éloignés conformément à la Loi 14 de l'IFAB.

Déclencheur de coup de pied de but : Un joueur doit être dans sa propre surface de but, en accord avec la Loi 16 de l'IFAB.

Déclencheur de corner : Un joueur doit être proche de l'un des coins du terrain, selon la Loi 17 de l'IFAB.

Déclencheur de touche : Un joueur doit se trouver au-delà de la ligne de touche, conformément à la Loi 15 de l'IFAB.

L'algorithme utilise des tolérances spécifiques pour chaque déclencheur afin de compenser les imprécisions potentielles dans les données de suivi.



Comment sont détectés les événements ?

Pour les phases arrêtées, l'algorithme distingue entre les moments où le ballon est en jeu et les moments où il est arrêté (balle morte). Un événement de balle morte se produit juste avant que le jeu ne s'arrête et est suivi par un événement de phase arrêtée pour reprendre le jeu. L'algorithme associe chaque événement de balle morte à un événement de reprise de jeu. Les hors-jeux sont traités comme des fautes car ils sont difficiles à différencier des autres infractions avec les données de suivi.

Les événements sont identifiés en vérifiant la position des joueurs et en s'assurant que le joueur qui doit reprendre le jeu se trouve à l'endroit approprié et que le ballon est près de lui. Si plusieurs joueurs peuvent reprendre le jeu, l'algorithme choisit celui qui est le plus proche du ballon.

Si les données montrent que le ballon passe la ligne de but à la fin d'une période, mais qu'il est difficile de dire si c'est un but, un événement "but ?" est ajouté pour indiquer l'incertitude.

Si un coup d'envoi n'est pas exécuté correctement et doit être répété, l'algorithme ajuste les informations pour refléter cette répétition.

Les tirs et les arrêts ?

Un tir est défini comme le moment où un joueur attaquant perd le ballon, qui peut ensuite mener à un but, un corner, un dégagement, ou un arrêt par le gardien. Les tirs bloqués ne sont pas considérés comme des tirs selon l'algorithme car ils impliquent une perte de possession suivie d'un gain par un autre joueur, et non par le gardien. Ces situations sont plutôt classées comme des passes.

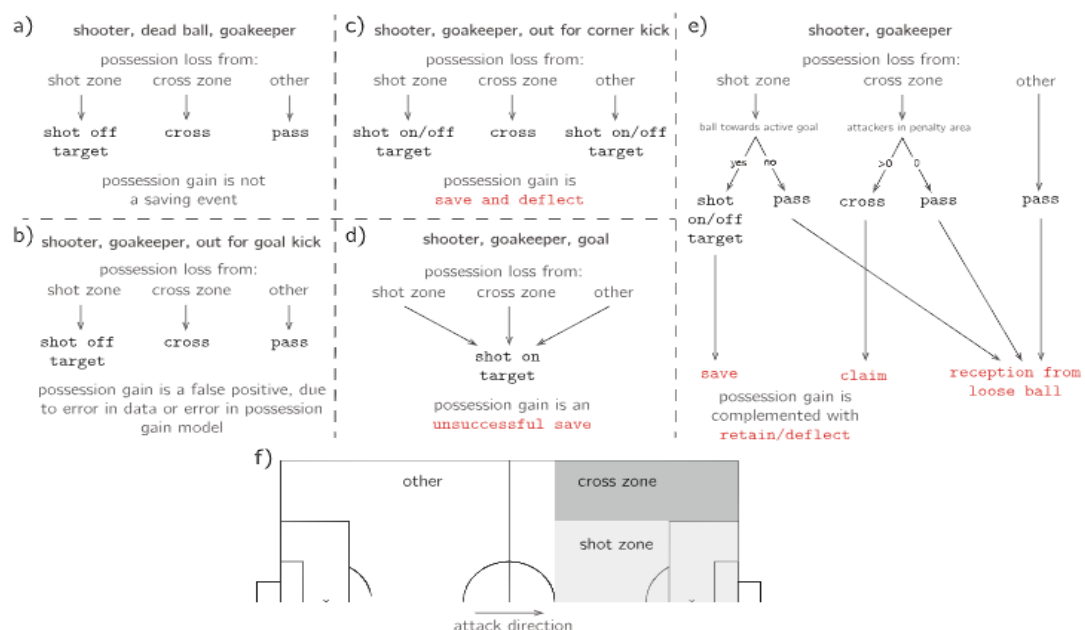
Un arrêt est défini comme un gain de possession par le gardien de but dans sa propre surface de réparation, précédé par un tir. L'algorithme examine également les cas où un ballon est dévié par un défenseur avant d'être capté par le gardien, mais ces situations sont souvent enregistrées comme des passes plutôt que des arrêts.

Les tirs sont classés en tirs cadrés (vers le but) ou hors cible, et les passes ou centres sont également distingués.

Les arrêts sont distingués entre les arrêts réussis (où le gardien garde le ballon ou le dévie efficacement) et les arrêts infructueux (où le ballon touche le gardien mais un but est quand même marqué).

Lors de l'analyse d'un tir ou d'un arrêt, l'algorithme prend en compte plusieurs facteurs tels que la position du tireur, la direction du ballon après la perte de possession, le nombre d'adversaires dans la surface de réparation, et si un intervalle de jeu arrêté se produit avant ou après que le gardien gagne la possession. Pour les arrêts, l'algorithme vérifie si le gardien perd la possession juste après l'arrêt pour différencier entre une rétention et une déviation du ballon.

La détermination si un tir est cadré ou non se fait en observant la trajectoire du ballon immédiatement après la perte de possession, avec une tolérance près des poteaux du but.



Autres

Un centre est défini par plusieurs critères, le ballon doit provenir de la zone de centre, le joueur suivant qui prend le contrôle du ballon (qu'il soit attaquant ou défenseur) doit se trouver dans la surface de réparation active. Au moins un joueur attaquant doit être présent dans la surface de réparation active. Si ces conditions sont remplies, la perte de possession est étiquetée comme un centre. Sinon, elle est considérée comme une passe.

Les gains de possession, en dehors des arrêts par les gardiens, sont classés comme des réceptions ou des interceptions. La distinction entre les deux dépend de la source de la perte de possession précédente, si la perte de possession provient d'un coéquipier, le gain est étiqueté comme une réception. Si la perte de possession provient d'un adversaire, il s'agit d'une interception.

Résultat ?

Des zones spécifiques autour des joueurs (50 cm ou 1 m selon le fournisseur de données) déterminent qui a la possession du ballon ou qui est en duel. Ces paramètres, appelés hyperparamètres, sont cruciaux pour la précision de l'algorithme.

L'algorithme a été testé contre des événements manuellement annotés pour vérifier sa précision.

Pour les événements comme les passes et les tirs, l'algorithme a réussi à identifier correctement la majorité des actions avec une précision et un rappel généralement supérieurs à 90% pour les passes.

Cependant, certains tirs ont été mal classés principalement à cause des limites des données de suivi, comme les occlusions (quand un joueur en cache un autre).

Les coups d'envoi et les penaltys ont été très bien identifiés. Les autres types d'événements comme les corners et les touches ont aussi montré une bonne précision, bien que quelques erreurs aient été notées à cause de la proximité de plusieurs joueurs ou de la mauvaise qualité des données de suivi.

Les données prédites par l'algorithme permettent de créer des cartes de chaleur montrant où les joueurs sont le plus souvent en possession du ballon, et comment ils passent, tirent ou interceptent le ballon.

On peut examiner comment les joueurs passent le ballon, où ils reçoivent les passes et comment ils avancent sur le terrain en comparant plusieurs matchs.