نشریه علمی مهندسی فناوری اطلاعات مکانی

سال دوازدهم، شماره سوم، پاییز ۱۴۰۳
Vol.12, No.3, Autumn 2024
۶۳ -۸۳
مقاله پژوهشی

خوشه‌بندی شکل تیم بر مبنای داده‌های مکانی-زمانی برای تجزیه و تحلیل رفتار جمعی بازیکنان تیم فوتبال

علی زارع زردینی*، زهرا بهرامیان

استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، دانشکدگان فنی، دانشگاه تهران

تاریخ دریافت مقاله: ۱۴۰۳/۰۴/۲۷ تاریخ پذیرش مقاله: ۱۴۰۳/۰۷/۱۷

چکیده

در سال‌های اخیر تحلیل داده‌های فوتبال برای آنالیز رفتار بازیکنان مورد توجه قرار گرفته است. بخش مهمی از این داده‌ها ماهیتی مکانی-زمانی دارند و همین امر اهمیت تحلیل‌های مکانی-زمانی را در صنعت فوتبال پررنگ‌تر می‌کند. هدف از این تحقیق، تحلیل رفتار جمعی بازیکنان در سطح ماکرو می‌باشد. برای این منظور، در گام اول در هر فریم زمانی، مشخصات تیم بر مبنای مجموعه‌ای از پارامترهای مکانی، هندسی، توپولوژیکی و توزیع استخراج می‌شود. سپس، این پارامترها مبنای خوشه‌بندی شکل تیم قرار می‌گیرند. این خوشه‌بندی به صورت دومرحله‌ای انجام می‌شود. در مرحله اول، خوشه‌های اصلی بر مبنای پارامترهای مکانی به دست می‌آید و وضعیت تدافعی یا تهاجمی بودن تیم مشخص می‌شود. در مرحله دوم به ازای هر یک از خوشه‌های اصلی، بر مبنای سایر پارامترها، خوشه‌های جدید تعریف می‌شوند که بیانگر شکل کلی تیم در یک محدوده مشخص می‌باشند. در این تحقیق از داده‌های یک مسابقه فوتبال استفاده شده و پنج خوشه اصلی، و همچنین پنج زیر خوشه به ازای هر یک از خوشه‌های اصلی به دست آمده است. در فرآیند ارزیابی، میزان تفاوت شکل تیم با مرکز خوشه متناظر با آن، اندازه‌گیری شده است. میزان انحراف معیار این تفاوت در خوشه‌های اصلی بین ۰٫۱۹ تا ۰٫۲۷ متغیر است. بر مبنای این تغییر در انحراف معیار، میزان نوسانات شکل تیم در مناطق مختلف زمین و بر مبنای خوشه‌بندی تیم در زمان‌های مختلف مسابقه، سهم زمانی خوشه‌ها و میزان تسلط تیم بر زمین، وضعیت تدافعی یا تهاجمی بودن و روند کلی تغییر وضعیت تیم مشخص می‌شود. در نظر گرفتن پارامترهای مکانی، توپولوژیکی و تراکم در کنار پارامترهای هندسی، انجام خوشه‌بندی به صورت دومرحله‌ای و بدون نیاز به انتقال شکل به فضای رستر (و در نتیجه عدم نیاز به استفاده از تکنیک‌های پردازش تصویر) از نقاط تمایز مقاله پیش‌رو در مقایسه با تحقیقات پیشین به شمار می‌آید.

کلیدواژه‌ها : تحلیل رفتار جمعی بازیکنان فوتبال، داده‌های مکانی-زمانی، پوش محدب، خوشه‌بندی دومرحله‌ای، الگوریتم *K-Means*

* نویسنده مکاتبه‌کننده: تهران، کارگرشمالی، دانشکدگان فنی، دانشگاه تهران، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی.
تلفن: ۰۹۱۲۴۵۲۷۸۵۹

Email: zare_zardiny@ut.ac.ir

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]

۱- مقدمه

امروزه صنعت فوتبال یکی از صنایع پول‌ساز در سراسر جهان به شمار می‌آید. با رشد فناوری در سال‌ها اخیر، این امکان فراهم شده است که در هر مسابقه فوتبال، طیف مختلفی از داده‌ها قابل جمع‌آوری باشد. فیلم مسابقات، موقعیت بازیکنان در زمین، آمار و نتایج مسابقات، رخدادهای بازی، داده‌های مرتبط با وضعیت سلامتی بازیکنان، آمار تماشاگران بازی، اطلاعات آماری بازی از جمله این داده‌ها به شمار می‌آیند که عموما حجم بالایی به خود اختصاص می‌دهند. این حجم و تنوع از داده‌ها در سال‌های اخیر توجه کارشناسان حوزه علم داده را به خود جلب نموده است و تحلیل این داده-ها در زمینه‌های مختلفی همچون ارزیابی عملکرد بازیکنان و تیم‌های فوتبال بیش از پیش مورد توجه قرار گرفته است [۱]. متناسب با نوع داده‌ها، طیف گسترده‌ای از تحلیل‌ها و پردازش‌ها قابل انجام می‌باشد. بخش قابل توجهی از این داده‌ها، ماهیتی کاملا مکانی-زمانی دارند [۲ و ۳]. از جمله این داده‌ها، می‌توان به موقعیت بازیکنان در زمین در هر لحظه از مسابقه، موقعیت توپ و همچنین مکان وقوع رخدادهایی که در جریان یک مسابقه فوتبال و در تعامل بین دو تیم رخ می‌دهد، اشاره کرد. تحلیل این داده‌های مکانی- زمانی می‌تواند در تصمیم‌گیری‌های مربیان نقش مهمی ایفا کند. با کمک این تحلیل‌ها، مربیان می‌توانند بینش و شناخت عمیق‌تری نسبت به الگوهای مکانی و موقعیتی که توسط تیم و حریف خود در طول بازی‌ها به نمایش گذاشته شده به دست آورند. این بینش می‌تواند در انتخاب ترکیب‌های ترجیحی، استراتژی‌های دفاعی و تهاجمی، موقعیت بازیکنان و الگوهای حرکتی به مربیان کمک کند [۴].

در یک نگاه کلی، می‌توان تحلیل‌های مکانی-زمانی داده‌های فوتبال را در دو سطح میکرو و ماکرو طبقه-بندی نمود [۵]. تحلیل‌های میکرو بر روی داده‌های مربوط به تک تک بازیکنان متمرکز می‌شوند و از آن‌ها برای ارزیابی رفتار و کارایی هر بازیکن در زمین استفاده می‌شود. میزان دوندگی، میزان مشارکت در رخدادهای مختلف، مسافت طی شده، سرعت، میزان اثرگذاری در رخدادهای بازی، کیفیت بازی بدون توپ از جمله این تحلیل‌ها به شمار می‌آید [۶، ۷ و ۸]. اما در سطح ماکرو، کل بازیکنان تیم در قالب یک مجموعه واحد مورد بررسی قرار می‌گیرند. به بیان مکانی، در این تحلیل‌ها قرار است به جای بررسی نقطه‌ای داده‌ها (برای تک تک بازیکنان)، یک هندسه واحد که معرف شکل تیم است به همراه مجموعه‌ای از خصوصیات استخراج شده برای کل تیم در نظر گرفته شود [۹]. این نوع از تحلیل‌ها بر روی حرکت تجمعی تیم و نحوه مدیریت فضا و زمان توسط کل مجموعه تیم متمرکز می‌شود [۱۰].

در این مقاله، از تحلیل‌های سطح ماکرو برای بررسی شکل و رفتار کل مجموعه تیم استفاده می‌شود. برای این منظور، در گام اول، هندسه و مشخصات کلی تیم در هر فریم زمانی از مسابقه، با استفاده از مجموعه‌ای از پارامترها شامل پارامترهای مکانی (مرکز هندسی شکل تیم، و مرکز هندسی بازیکنان)، هندسی (مساحت و محیط شکل، طول و عرض محدوده، میزان مستطیلی بودن شکل، میزان دایره‌ای بودن شکل، میزان فشردگی شکل، و تعداد رئوس شکل)، توپولوژیکی (مساحت محدوده مشترک با زون‌های زمین، و مالکیت توپ)، و پارامترهای تراکم و توزیع استخراج می‌گردد. سپس در گام دوم، بر مبنای این پارامترها شکل تیم در طول مسابقه با استفاده از روش *K-Means* در دو مرحله خوشه‌بندی می‌شود. بدین ترتیب شکل کلی تیم در زمان‌های مختلف بازی با استفاده از خوشه‌های اصلی و همچنین خوشه‌های فرعی به ازای هر یک از خوشه‌های اصلی مشخص می‌گردد. با بررسی ارتباط زمانی و مکانی خوشه‌ها با محدوده زمین و رخدادهای مسابقه، رفتار کلی تیم شامل میزان حضور و تسلط تیم بر زمین، وضعیت تدافعی یا تهاجمی بودن تیم، نحوه تعامل تیم در وضعیت‌های مختلف بازی و نیز روند کلی تغییر جریان حرکت تیم قابل تفسیر می‌باشد.

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]

در تحقیقات پیشین، سعی شده است با روش‌های مختلف رفتار کلی بازیکنان تیم مورد بررسی قرار گیرد. در اکثر این تحقیقات، تمرکز صرفا بر روی موقعیت تک‌تک بازیکنان بوده است و شکل واحدی برای تیم در نظر گرفته نشده است. در تحقیقاتی نیز که تحلیل‌های خود را بر مبنای شکل تیم انجام داده‌اند، یا صرفا خصوصیات عددی (مانند مساحت، طول و عرض و ...) را مد نظر قرار داده‌اند و یا صرفا از دید هندسی (مستقل از موقعیت و ابعاد) شکل تیم را در طول مسابقه مورد ارزیابی قرار داده‌اند. مقاله پیش رو با توجه به این کاستی‌ها و با در نظر گرفتن خصوصیات مختلف شکل تیم، روشی برای تحلیل رفتار کلی تیم ارائه داده است. در نظرگرفتن پارامترهای مکانی، هندسی، توپولوژیکی و تراکم، انجام فرآیند خوشه‌بندی به صورت دو مرحله‌ای، همچنین انجام تحلیل‌ها بدون نیاز به انتقال شکل تیم به فضای رستر (و در نتیجه عدم نیاز به استفاده از تکنیک‌های پردازش تصویر) از نقاط تمایز مقاله پیش‌رو به شمار می‌آید.

تحقیق پیش رو در پنج بخش تنظیم شده است. بخش دوم به مرور مطالعات و پژوهش‌های پیشین در زمینه مدلسازی و تحلیل شکل تیم می‌پردازد. در بخش سوم، توضیحاتی در ارتباط با داده‌های مکانی در یک مسابقه فوتبال و آنالیز شکل تیم بر مبنای این داده‌ها ارائه می‌شود. در بخش چهارم، روش پیشنهادی برای تحلیل رفتار جمعی بازیکنان تیم فوتبال تشریح می‌گردد. در بخش پنجم نتایج پیاده‌سازی روش پیشنهادی نشان داده شده و از بعد مکانی تفسیر می‌شود. در نهایت، بخش ششم به نتیجه‌گیری و ارائه پیشنهادهایی برای مطالعات آتی می‌پردازد.

۲- پیشینه تحقیق

با فراهم شدن امکان جمع‌آوری داده‌های فوتبال و مشخص شدن اهمیت تحلیل داده‌ها بر تصمیم‌گیری‌ها و برنامه‌ریزی‌های مربیان و باشگاه‌های فوتبال، موضوع تحلیل داده‌های مکانی-زمانی فوتبال بیش از پیش مورد توجه پژوهشگران قرار گرفته است. این پژوهش‌ها هم در سطح میکرو و با بررسی رفتار تک تک بازیکنان و هم در سطح ماکرو و با تمرکز بر رفتار کلی تیم انجام شده است. در این بخش به مرور تحقیقات در حوزه تحلیل رفتار و شکل کلی تیم پرداخته می‌شود.

ژانگ[1] (۲۰۲۲) رفتار تکنیکال تیم در موقعیت‌های حمله، دفاع، و انتقال بازی از حمله به دفاع و از دفاع به حمله بر مبنای داده‌های موقعیت بازیکنان را مورد ارزیابی قرار داده است [۱۰]. برای این منظور، ابتدا وضعیت تیم بر مبنای داده‌های رخدادهای بازی مشخص شده است. سپس مرکز هندسی، طول، عرض و مساحت تیم، و میزان جابه‌جایی تیم در فازهای چهارگانه محاسبه شده است و در طول زمان بر مبنای پارامترهای آماری میانگین و واریانس باهم مقایسه شده‌اند. شکلی که در این مقاله برای تعریف هندسه تیم در نظر گرفته شده است، یک پوش‌محدب می‌باشد. مطابق با نتایج به دست آمده رفتار تاکتیکی تیم به طور قابل توجهی در فاز مالکیت توپ با فازهای دیگر مسابقه متفاوت است و تیم مساحت بیشتری از سطح زمین را پوشش می‌دهد. در زمان تغییر فاز بین عدم مالکیت به مالکیت عرض تیم در مقایسه با طول تیم افزایش بیشتری پیدا می‌کند و همچنین در زمان انتقال بازی از فاز حمله به دفاع گستردگی بیشتری در مقایسه با زمان از دست توپ برای تیم مشاهده می‌شود.

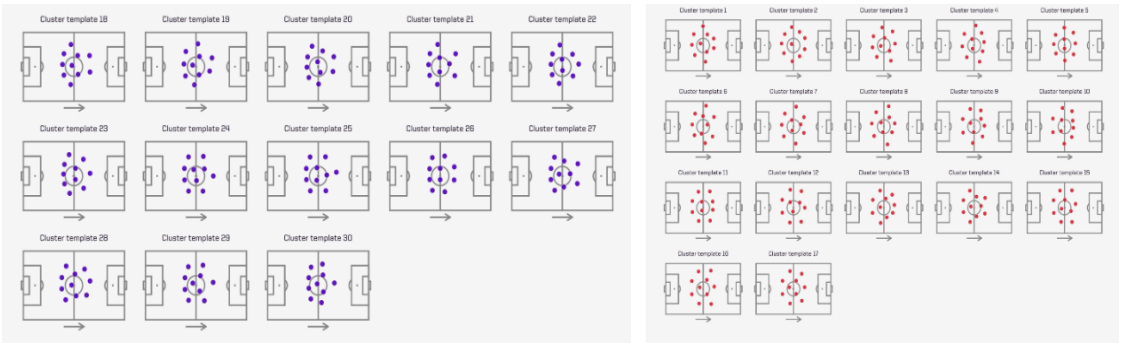
شاو[2] و همکارش (۲۰۲۰) بر مبنای موقعیت بازیکنان، روشی برای تعیین ترکیب تدافعی و تهاجمی تیم در طول مسابقه ارائه داده است [۱۱]. این روش مبتنی بر خوشه‌بندی سلسله مراتبی تجمعی داده‌های موقعیت بازیکنان است. در این محاسبات، داده‌ها در بازه‌های زمانی ۲ ثانیه‌ای تجمیع و سپس تحلیل شده‌اند. در نهایت برای برآورد احتمال قرار گیری بازیکنان در فرمت‌های شناسایی شده، از الگوریتم انتخاب مدل بیزی استفاده شده است.

[1] *Zhang*
[2] *Shaw*

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]
[DOI: 10.61186/jgit.12.3.63]

بوئنو و همکارانش (۲۰۲۱) برای بررسی شکل تیم از روش منحنی فراکتال چند مقیاسی (*Multiscale Fractal Curve*)، به عنوان یک روش تحلیل شکل مبنی بر ناحیه استفاده کرده‌اند [۹]. در این روش، شکل تیم با یک پوش محدب تعریف شده است و سپس این شکل برای انجام تحلیل‌های بعدی به یک تصویر رستری تبدیل شده است. آنگاه بر مبنای مشخصات پیکسلی تصاویر مورد، فرآیند خوشه‌بندی با روش *K-Means* انجام شده است. تمرکز اصلی این مقاله بر روی بعد هندسی شکل بوده است و پارامترهای مربوط به موقعیت قرارگیری تیم در آن لحاظ نشده است. همچنین در این مقاله به دلیل استفاده از منحنی فراکتال چند مقیاسی که روشی مستقل از مقیاس، دوران، و جابه‌جایی است، نمی‌تواند تحلیلی از کیفیت بازی تیم ارائه دهد و صرفا به خوشه‌بندی هندسی شکل تیم بسنده می‌کند.

ویتمور و همکارش (۲۰۲۱) با توجه به اینکه ترکیب قرارگیری بازیکنان در طول مسابقه می‌تواند متغییر باشد، به بررسی وضعیت تیم در دو حالت مالکیت یا عدم مالکیت توپ پرداخته و محتمل‌ترین شکل وضعیت تیم را شناسایی کرده‌اند [۱۲]. برای این منظور، در فواصل زمانی متمایز، میانگین موقعیت بازیکنان محاسبه شده و بر مبنای آن فرآیند خوشه‌بندی انجام شده است. بر مبنای نتایج خوشه‌بندی، محتمل‌ترین شکل تیم را از الگوهای ارائه شده در شکل (۱) برای دو وضعیت مالکیت یا عدم مالکیت توپ انتخاب نموده است.



(ب) (الف)

شکل۱: (الف) الگوهای کلی تیم در وضعیت مالکیت توپ، (ب) الگوی کلی تیم در وضعیت عدم مالکیت توپ [۱۲]

گوئز[1] و همکارانش (۲۰۲۰) در مرور تحقیقات انجام شده در ارتباط با استفاده از داده‌های موقعیت بازیکنان در پشتیبانی از تجزیه و تحلیل عملکرد تاکتیکی به اهمیت توابع مکانی تجمعی اشاره کرده‌اند [۱]. توابع مکانی تجمعی ضمن مدلسازی رفتار بازیکنان در سطح ماکرو در هر بازه زمانی، پیچیدگی تحلیل‌ها را تا سطح قابل تفسیر کاهش می‌دهد. پارامترهای مانند مرکز هندسی نقطه‌ای، مرکز هندسی خطی، میانگین فاصله بازیکنان تا مرکز هندسی تیم، سطح تحت پوشش محدوده تیم، و طول و عرض محدوده تیم از جمله پارامترهایی است که در این مقاله به آن‌ها اشاره شده است.

گونسالوز[2] و همکارانش (۲۰۱۹) با توجه به تغییرات زمانی ترکیب بازیکنان یک تیم در زمان حمله و دفاع، روش سلسله مراتبی و آنالیز سری‌های زمانی را برای تعیین ترکیب و چینش بازیکنان استفاده کرده‌اند [۱۳].

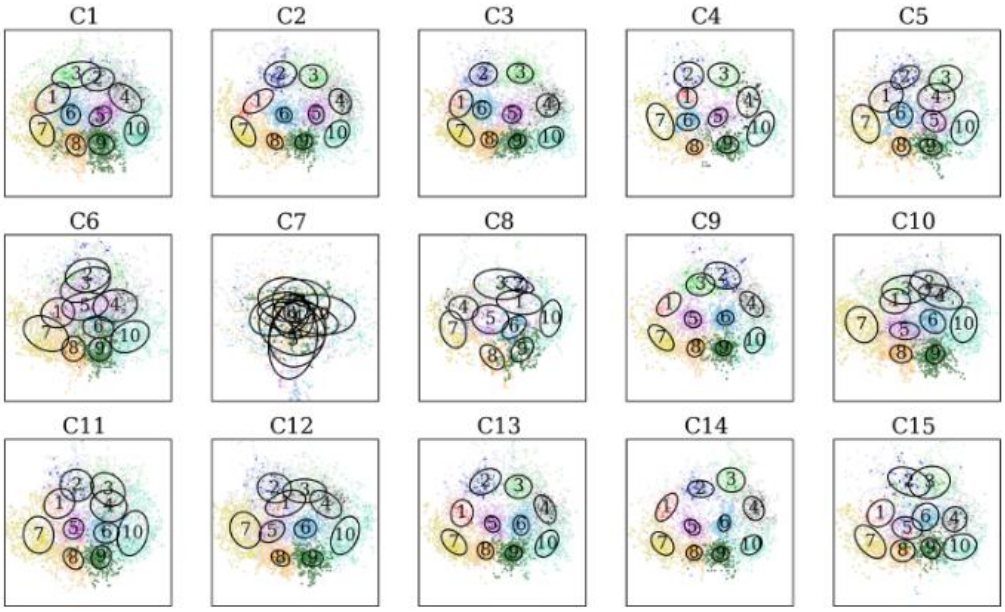
[1] *Goes*

[2] *Goncalves*

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]

این خوشه‌بندی بر مبنای موقعیت بازیکنان و مساحت تحت تسلط هر یک از آن‌ها که با دیاگرام ورونوی استخراج شده است و بر مبنای چند مسابقه برگزار شده توسط یک تیم انجام شده است. در شکل (۲) نمایی از خروجی این خوشه‌بندی نمایش داده شده است.



شکل۲: خوشه‌بندی بازیکنان بر مبنای موقعیت و مساحت محدوده تحت تسلط [۱۳]

نارizوکا[1] و همکارش (۲۰۱۹) برای تحلیل مکانی-زمانی رفتار جمعی تیم، فضای موثر بازی یک تیم را در قالب یک پوش محدب تعریف کرده‌اند و در کنار آن خصوصیاتی چون مالکیت توپ، مسافت طی شده، نسبت تغییرات طول و عرض محدوده موثر تیم، مالکیت توپ، محدوده حد فاصل موقعیت پیشگام‌ترین بازیکن تا خط دروازه را مورد بررسی قرار داده‌اند [۱۴]. سپس با مقایسه این پارامترها سعی شده است که موقعیت نسبی تیم نسبت به تیم‌های صدر و قعر جدول امتیازبندی و همچنین اثرگذاری پارامترهای مختلف بر نتایج، مورد ارزیابی قرار گیرد. مطابق با نتایج به دست آمده، مالکیت توپ بیشترین تاثیر را بر الگوهای رفتار جمعی تیم داشته است.

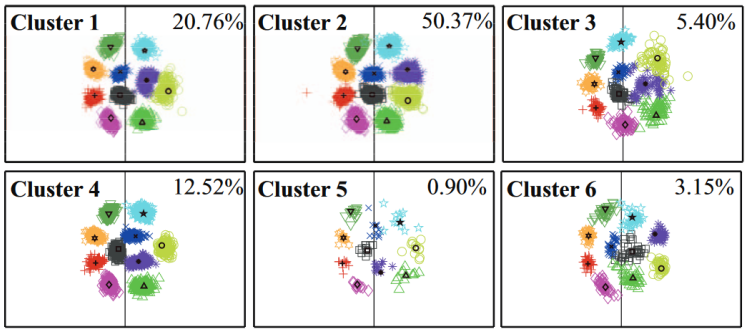
بیالکوفسکی[2] و همکارانش (۲۰۱۴) به منظور تجزیه و تحلیل بزرگ مقیاس مسابقات فوتبال، داده های مکانی-زمانی موقعیت بازیکنان را مورد استفاده قرار داده‌اند [۱۵]. برای این منظور بر روی داده‌ها فرایند خوشه‌بندی به روش *K-Means* انجام شده است و علاوه بر موقعیت، نقش بازیکنان نیز در نظر گرفته شده است. خروجی این فرآیند شناسایی ترکیب بازیکنان در طول بازی و در وضعیت‌های مختلف بازی بوده است. شکل (۳) نمایی از خروجی این خوشه‌بندی را نشان می‌دهد.

[1] *Narizuka*

[2] *Bialkowski*

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]



شکل ۳: خوشه‌بندی موقعیت بازیکنان به روش *K-Means* [۱۵]

در یک جمع‌بندی کلی می‌توان گفت که در اکثر تحقیقات انجام شده، تمرکز پردازش‌ها از بعد مکانی بر روی داده‌های نقطه‌ای موقعیت بازیکنان بوده است و کمتر شکل کلی تیم در نظر گرفته شده است [۱۱، ۱۲، ۱۳، ۱۴ و ۱۵]. در این مقالات با خوشه‌بندی موقعیت مکانی-زمانی بازیکنان، آرایش کلی بازیکنان در طول مسابقه مشخص شده است، اما این شیوه نمایش ساختار تیم، حس ملموسی از میزان تسلط تیم بر زمین مسابقه و تغییرات کلی شکل تیم در طول زمان را ارائه نمی‌دهد. بر همین اساس در مقالات دیگر تمرکز اصلی تحلیل‌ها بر روی هندسه واحدی بوده است که در هر فریم زمانی برای کل تیم در نظر گرفته شده است [۱ و ۱۰]. هرچند که نتایج این مقالات با این رویکرد، به تحلیل‌های سطح ماکرو نزدیک‌تر است، اما در اینجا نیز کاستی‌های مهمی به چشم می‌خورد. در این مقالات سعی شده است که شکل تیم صرفا بر مبنای یک سری خصوصیت عددی همچون مساحت شکل تیم، طول و عرض تیم و تغییرات این مقادیر در طول زمان مورد بررسی قرار گیرد؛ در حالی که کمتر به ماهیت هندسی شکل تیم پرداخته شده است. در این بین، بوئنو و همکاران (۲۰۲۱) فرآیند خوشه‌بندی را بر مبنای شکل هندسی تیم انجام داده و در آن خروجی‌های ملموسی از این خوشه‌بندی شکل تیم ارائه داده‌اند [۹]. اما در اینجا نیز این خوشه‌بندی، صرفا هندسی بوده است و مستقل از مساحت، دوران تیم و حتی موقعیت کلی تیم در زمین مسابقه انجام شده است. این که یک تیم در کدام محدوده از زمین، با چه شکل هندسی، در کدام امتداد و با چه مساحت تحت تسلطی در حال بازی است موضوعی است که در آنالیز رفتار تیم بسیار حائز اهمیت می‌باشد اما همه این پارامترها، کمتر به صورت یکجا در یک مقاله مورد بررسی قرار گرفته‌اند. بر همین اساس در مقاله پیش رو سعی شده است تا با توجه به همه این کاستی‌ها، روشی برای تحلیل کلی رفتار تیم ارائه شود. تمرکز بر شکل کلی تیم به جای موقعیت تک تک بازیکنان، در نظر گرفتن پارامترهای هندسی توصیف کننده شکل تیم در کنار پارامترهای عددی و همچنین پارامترهای توپولوژیکی و تراکم، انجام خوشه-بندی تیم در دو فاز (بر مبنای موقعیت تیم و بر مبنای هندسه و سایر مشخصات تیم)، انجام پردازش‌ها بدون نیاز به انتقال شکل تیم به فضای رستر (و بدون نیاز به تکنیک‌های پردازش تصویر) از نقاط تمایز این طرح، در مقایسه با پژوهش‌های پیشین به شمار می‌آید.

۳- آنالیز شکل تیم بر مبنای داده‌های مکانی در فوتبال

به طور کلی داده‌های مکانی فوتبال شامل سه بخش موقعیت‌های مکانی-زمانی بازیکنان در محدوده زمین، موقعیت توپ در هر لحظه از مسابقه و موقعیت رخداد-های بازی می‌باشند [۱۶]. امروزه عموما موقعیت بازیکنان و توپ با کمک سیستم‌های ردیابی ماهواره‌ای مانند *GPS* جمع‌آوری می‌شود [۱۰]. با وجود این، روش‌های دیگری همچون سیستم‌های ردیابی اپتیکال (با کمک پردازش تصویر و فیلم)، و سیستم‌های ردیابی

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]
[DOI: 10.61186/jgit.12.3.63]

رادیویی (با کمک ارسال سیگنال‌های الکتریکی) می‌توانند به جمع‌آوری این داده‌ها کمک کنند. داده‌های مربوط به رخدادهای بازی معمولا توسط افرادی که فیلم مسابقه را در نرم‌افزارهای خاص مشاهده می‌کنند، جمع آوری می‌شوند.

بر مبنای ماهیت رخدادها، طبقه‌بندی‌های مختلفی برای آن‌ها در نظر گرفته می‌شود. رخدادها می‌توانند در یک نقطه مشخص رخ دهند (مانند رخداد وقوع خطا) و یا یک مبدا و مقصد مشخص داشته باشد (مانند رخداد پاس). از طرف دیگر، رخدادها می‌توانند ناظر بر فعالیت بازیکنان باشد (مانند رخدادهای پاس و شوت) و یا بیان‌کننده یک رخداد فنی باشند (مانند رخدادهای وقوع خطا، آغاز و پایان نیمه‌های بازی) [۳]. همچنین در یک طبقه‌بندی دیگر می‌توان آن‌ها را در دو بخش رخدادهای مبتنی بر توپ (مانند رخداد پرتاب اوت، پاس، پاس در عمق، شوت، و دریافت توپ)، و رخداد-های بدون توپ (مانند تکل، و دریبل) مورد بررسی قرار داد [۱۷].

منابع مختلفی برای دسترسی به داده‌های مکانی فوتبال وجود دارد. این داده‌ها معمولا با فرمت *CSV* یا *JSON* برای مسابقات مختلف در دسترس هستند. داده-های رخدادهای بازی می‌تواند حجمی در حدود چند صد کیلوبایت داشته باشند در حالی که حجم داده‌های موقعیت بازیکنان ممکن است به چندین مگابایت نیز برسد. هرچند تامین‌کنندگان متعددی برای جمع‌آوری داده وجود دارند که در ازای هزینه چنین داده‌های را در اختیار قرار می‌دهند، اما منابع رایگانی نیز برای دسترسی و دانلود این داده‌ها وجود دارد. از جمله مهم-ترین این منابع می‌توان به داده ها استاسبمب[1][۱۸]، کگل[2][۱۹]، ویسکوت[3][۲۰]، داده‌های نمونه متریکا اسپرت[4][۲۱] و افبیرف[5][۲۲] اشاره کرد.

با جمع‌آوری داده‌های مکانی از یک مسابقه فوتبال می-توان نسبت به تحلیل‌های مختلف به هدف آنالیز رفتار تیم و کارایی بازیکنان اقدام نمود. یکی از انواع این آنالیزها، آنالیز تکنیکال می‌باشد. آنالیز تکنیال در فوتبال موضوعی است که بر روی حرکت تجمعی تیم و نحوه مدیریت فضا و زمان توسط کل مجموعه تیم متمرکز می‌شود [۱۰]. در نخستین گام برای انجام تحلیل داده‌ها در سطح کل تیم، لازم است برای هر فریم زمانی از مسابقه، یک شکل کلی برای تیم در نظر گرفته شود. این شکل ساختاردهی کلی تیم در سطح زمین بازی را مشخص می کند. به طور کلی بر مبنای پژوهش‌های انجام شده می‌توان برای تعریف شکل کلی تیم راهکارهای مختلف زیر را در مد نظر قرار داد:

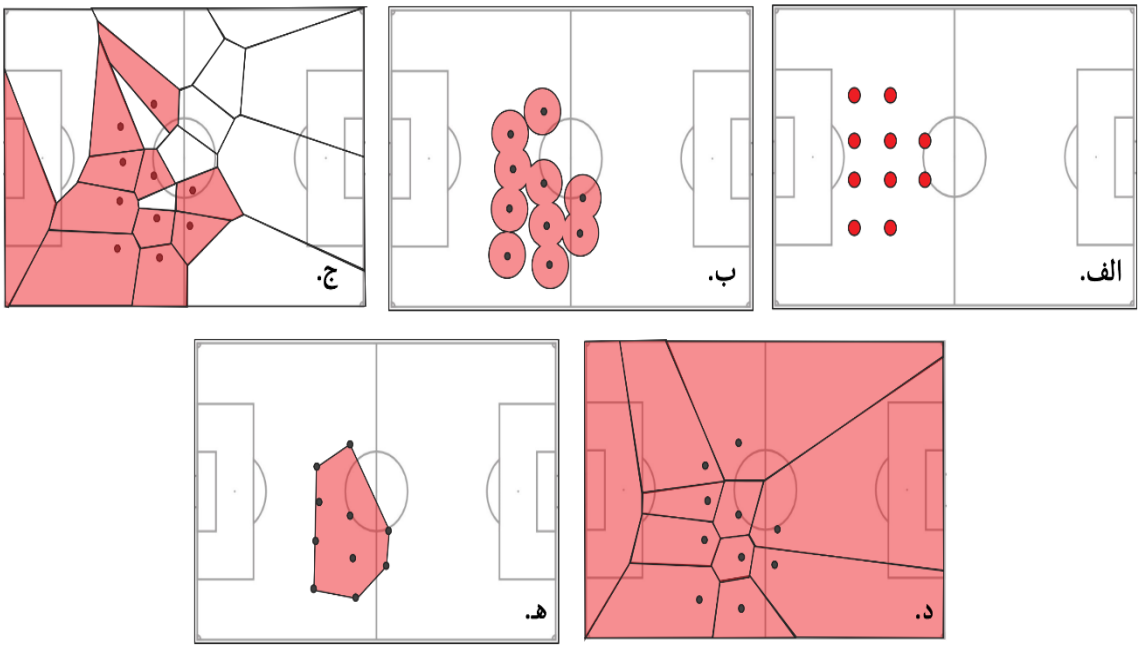
- ترکیب‌های معمول بازیکنان در یک مسابقه: مرسوم-ترین الگویی که برای بیان شکل تیم ارائه می‌شود، فرمت‌های قرارگیری آغازین بازیکنان بر مبنای موقعیت متوسط آن‌ها می باشد [۱۱] (شکل ۴-الف). یک تیم می‌تواند با ترکیب مشخص، بازی را آغاز کند، اما این ترکیب پویا بوده و می‌تواند در طول مسابقه مدام تغییر شکل دهد. همچنین از لحاظ بصری، اطلاعاتی از میزان پوشش و تسلط تیم بر زمین بازی ارائه نمی‌دهد.
- در نظر گرفتن حریم پیرامون بازیکنان تا شعاع مشخص: می‌توان شعاع فعالیت هر بازیکن را با یک حریم دایره‌ای مشخص نمود و اجتماع این دوایر را به عنوان محدوده فعالیت کل تیم در نظر گرفت (شکل۴-ب). این شعاع‌ها می‌تواند بر مبنای رخداد-های مرتبط با بازیکن مانند دریبل، دفاع و پاس تعیین شوند [۱۶]. تغییر مداوم موقعیت بازیکنان در طول بازی می‌تواند منجر به تغییر تحدب شکل و همچنین ایجاد حفره در شکل کلی شود و این

[1] *StatsBomb Open Data*
[2] *Kaggle*
[3] *Wyscout*
[4] *Metrica Sports Sample Data*
[5] *FBref*

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]
[DOI: 10.61186/jgit.12.3.63]

موضوع می‌تواند پیچیدگی محاسبات بر روی شکل تیم را به همراه داشته باشد.

- دیاگرام ورونوی بازیکنان تیم: با در نظر گرفتن موقعیت بازیکنان در هر زمان و ترسیم دیاگرام ورونوی آن‌ها می‌توان کل زمین را به محدوده‌های تحت پوشش هر بازیکن تیم افراز نمود [۲۳]. برای ترسیم دیاگرام ورونوی لازم است که موقعیت بازیکنان هر دو تیم در نظر گرفته شود که در این حالت شکل نهایی شکل گسسته‌ای می‌شود (شکل۴-ج) و یا صرفا بر مبنای موقعیت بازیکنان یک تیم ساخته شود که در این حالت محدوده کلی شکل تیم با محدوده کل زمین یکسان می‌شود (شکل ۴-د).
- پوش محدب: این شکل محدوده‌ای که توسط تیم در هر لحظه از بازی پوشش داده می‌شود را نشان می‌دهد [۱۶] و در [۱۳] این محدوده به عنوان محدوده موثر بازی معرفی شده است. این شکل به عنوان تابعی از زمان و موقعیت تمامی بازیکنان (به جز دروازه‌بان) در آن زمان مشخص ساخته می‌شود (شکل ۴-هـ). این شکل یکپارچه و محدب بوده و بینشی بصری از میزان تسلط تیم بر بازی را ایجاد می‌کند.



شکل۴: اشکال مختلف تیم، (الف) ترکیب آغازین بازیکنان، (ب) دوایر پیرامون بایکنان، (ج) دیاگرام ورونوی گسسته، (د) دیاگرام ورونوی پیوسته، (هـ) پوش محدب

یکی از تحلیل‌هایی که می‌توان در سطح ماکرو برای تحلیل رفتار تیم مورد استفاده قرار داد، خوشه‌بندی هندسه تیم در زمان‌های مختلف مسابقه می‌باشد. اینکه یک تیم در چه زمان‌هایی از مسابقه چه شکلی به خود می‌گیرد می‌تواند به مربیان در شناسایی نقاط ضعف و قوت تیم در مواجه با تیم مقابل کمک کند. برای خوشه‌بندی به مجموعه‌ای از پارامترها نیاز است که بتوانند شکل تیم را در هر فریم زمانی توصیف کنند. در یک دید کلی، می‌توان روش‌های توصیف هندسی شکل‌ها را در سه بخش اصلی طبقه‌بندی کرد:

- روش‌های مبتنی بر کانتور: این روش‌ها یک سری از مشخصات شکل اصلی را به یک نمایش عددی تبدیل می‌کنند، به گونه‌ای که ویژگی‌های هندسی جزئی شکل را حفظ می‌کند. محاسبات در این روش عموما

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]
[DOI: 10.61186/jgit.12.3.63]

در فضای برداری انجام می‌شود و عموما از مرز چندضلعی برای توصیف شکل استفاده می‌شود. از آنجایی که روش‌های مبتنی بر کانتور بر روابط متقابل بین رئوس تأکید دارند، مزیت کارآمدی محاسباتی را دارند [۲۴].

- روش‌های مبتنی بر یادگیری عمیق که در واقع یادگیری با سطوح مختلف بازنمایی است. در این روش‌ها، ورودی بدون ساختار به یک نمایش خروجی ایده‌آل مانند گراف تبدیل می‌شود و محاسبات بر روی این خروجی انجام می‌شود [۲۴].
- روش‌های مبتنی بر ناحیه که با انتقال شکل از فضای برداری به فضای رستری سعی در انعکاس محتوی داخلی شکل دارند. در این فضا، توصیف شکل و انجام محاسبات بر مبنای پیکسل‌ها انجام می‌شود [۹].

در این مقاله از روش‌های مبتنی بر کانتور برای توصیف شکل استفاده می شود. چرا که بر خلاف روش‌های یادگیری ماشین به حجم بالایی از داده‌ها و محاسبات نیاز ندارند. همچنین در روش‌های مبتنی بر ناحیه، لازم است یکبار شکل مورد نظر به یک تصویر تبدیل شود و سپس محاسبات بر روی آن انجام شود اما در روش‌های مبتنی بر کانتور می‌توان توصیف‌کننده عددی متعددی را مستقیما از روی شکل تعریف نمود. نکته مهمی که اشاره به آن لازم است آنکه شکل ممکن است به طور کامل از توصیف‌گرها قابل بازسازی نباشد، اما توصیف گرهایی که از روی اشکال تعریف می‌شوند باید به اندازه کافی متفاوت باشند که بتوان بر مبنای آن اشکال را از یکدیگر متمایز کرد [۲۵].

۴- روش پیشنهادی

در این مقاله، به تحلیل تکنیکال شکل کلی تیم در سطح ماکرو با استفاده از داده‌های فوتبال پرداخته می‌شود که این تحلیل در سطح کل تیم صورت می‌گیرد. برای این منظور، برای هر فریم زمانی از مسابقه، یک شکل کلی برای تیم در نظر گرفته می‌شود. با توجه به یکپارچگی شکل ایجاد شده توسط پوش محدب و محدب بودن آن و همچنین بینش بصری که این شکل از تسلط تیم بر زمین بازی ایجاد می‌کند، از پوش محدب به عنوان شکل کلی تیم استفاده می‌شود.

برای تحلیل رفتار تیم، هندسه تیم در زمان‌های مختلف مسابقه خوشه‌بندی می‌شود. برای این منظور، از روش‌های مبتنی بر کانتور برای توصیف شکل استفاده می‌شود؛ چرا که بر خلاف روش‌های یادگیری ماشین به حجم بالایی از داده‌ها و محاسبات نیاز ندارند. همچنین در روش‌های مبتنی بر ناحیه، لازم است یک بار شکل مورد نظر به یک تصویر تبدیل شود و سپس محاسبات بر روی آن انجام شود، اما در روش‌های مبتنی بر کانتور می‌توان توصیف‌کننده عددی متعددی را مستقیما از روی شکل تعریف نمود. روش پیشنهادی در دو گام شامل استخراج پارامترهای خوشه‌بندی و اعمال روش خوشه‌بندی انجام می‌شود. در ادامه این دو گام تشریح می‌گردند. روال کلی روش پیشنهادی مطابق با شکل(۵) می‌باشد.

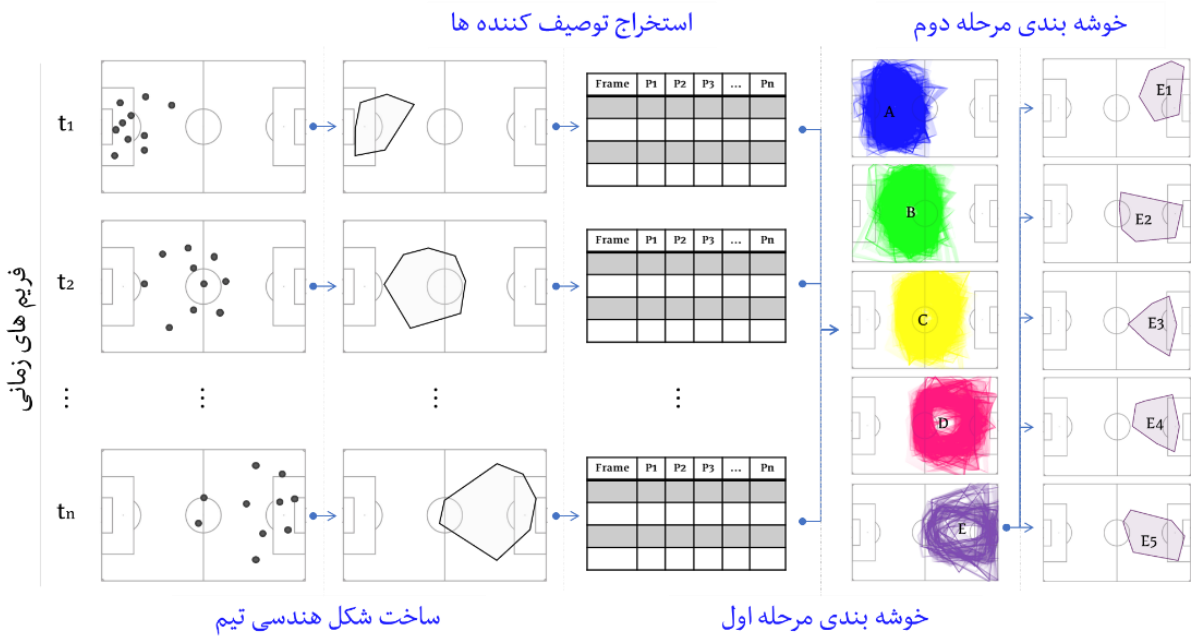
۴-۱- گام اول: استخراج پارامترهای خوشه‌بندی

در این مقاله، ابتدا از داده‌های موقعیت بازیکنان، داده‌های خارج از زمان قانونی بازی حذف می‌شوند و موقعیت بازیکنان تعویض شده جایگزین می‌گردند. علاوه بر این، داده‌های هر دو نیمه بازی به یک سمت انتقال می‌یابند. سپس در هر فریم زمانی، پوش محدب برمبنای موقعیت بازیکنان ساخته می‌شود. از طرف دیگر، از میان داده‌های موقعیت توپ، داده‌های خارج از محدوده زمین حذف می‌گردند.

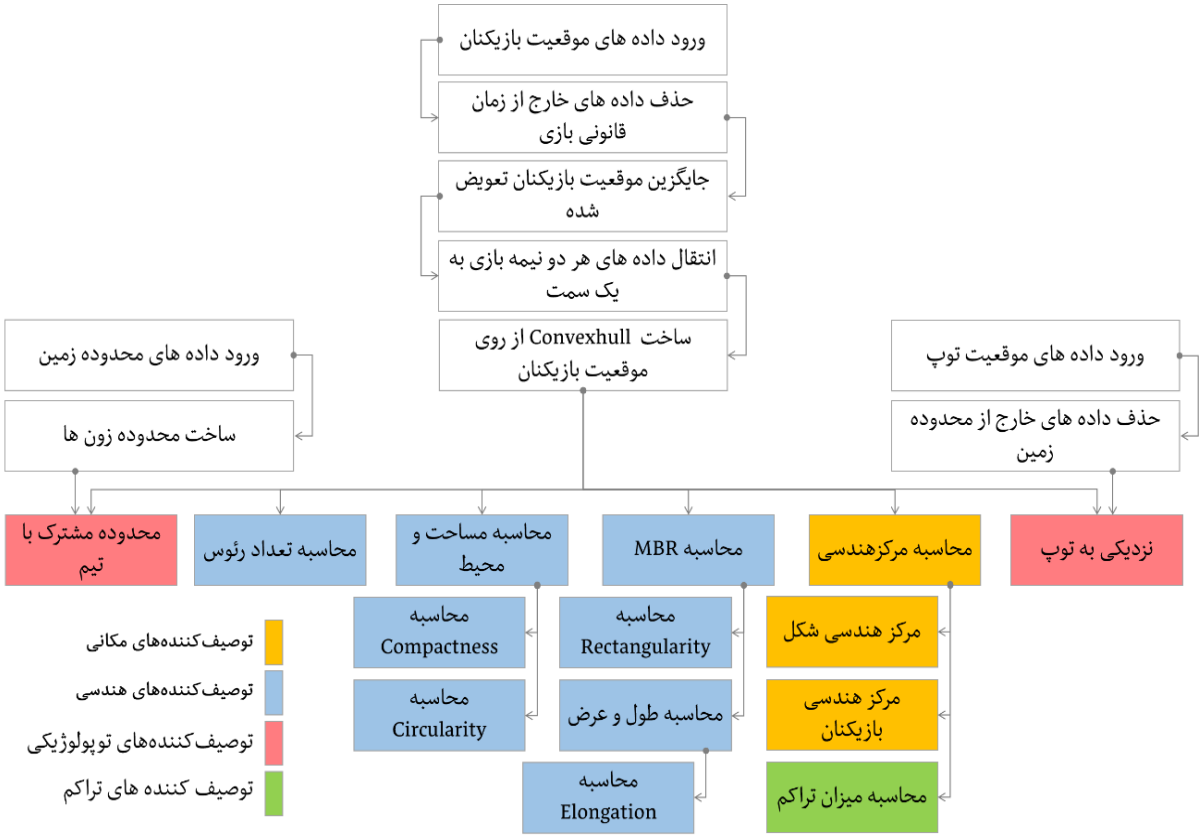
سپس پارامترهایی که برای توصیف شکل در فضای برداری استفاده می شوند در چهار طبقه کلی شامل توصیف‌کننده‌های هندسی، توصیف‌کننده‌های مکانی، توصیف‌کننده‌های توپولوژیکی، و توصیف‌کننده‌های تراکم و توزیع محاسبه می‌شوند. در این مقاله، پارامتر-های توصیف‌کننده شکل، مطابق با روند توضیح داده شده در شکل (۶) برای هر فریم زمانی محاسبه می-گردند.

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]



شکل۵: روال کلی در روش پیشنهادی



شکل۶: محاسبه پارامترهای توصیف‌کننده شکل

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]
[DOI: 10.61186/jgit.12.3.63]

در ادامه، به ارائه توضیحات در مورد هر یک از این توصیف‌کننده‌ها پرداخته می‌شود.

• **توصیف‌کننده‌های مکانی:** این پارامترها که بیان‌کننده موقعیت کلی تیم در محدوده زمین مسابقه هستند و عبارتند از:

o مرکز هندسی شکل تیم: این مرکز صرفا بر مبنای محدوده پوش محدب و مختصات رئوس آن طبق رابطه (۱) محاسبه می‌شود. در این رابطه N تعداد رئوس پوش محدب و Vi مختصات رئوس می باشد.

رابطه(۱)

$$\text{Convexhull Centroid}=\frac{1}{N}\sum_{i=1}^{N}V_i$$

o مرکز هندسی بازیکنان: این مرکز بر مبنای موقعیت تمام بازیکنان موجود در زمین به جز دروازه بان و مطابق با رابطه(۲) به دست می‌آید. در این رابطه Pi مختصات بازیکنان می باشد.

رابطه(۲)

$$\text{Players Centroid}=\frac{1}{10}\sum_{i=1}^{N}V_i$$

• **توصیف‌کننده‌های هندسی:** مهم‌ترین پارامترهای هندسی که برای توصیف شکل تیم مورد استفاده قرار می‌گیرند عبارتند از:

o مساحت و محیط شکل: متریک قابل قبولی از میزان پوشش محدوده زمین توسط تیم و میزان اشراف کلی بر زمین بازی بیان می‌کنند.

o طول (L) و عرض (W) محدوده: با کمک طول (L) و عرض (W) محدوده توصیف بهتری از میزان کشیدگی شکل به دست می‌آید [۲۶]. سپس میزان کشیدگی[1] شکل بر مبنای رابطه (۳) تعریف می‌شود:

رابطه(۳)

$$\text{Elongation} = 1-\frac{W}{L}$$

o میزان مستطیلی بودن شکل:[2] میزان نزدیکی یک شکل به کوچک‌ترین مستطیل محاط آن بر مبنای رابطه (۴) محاسبه می‌شود [۲۶].

رابطه(۴)

$$\text{Rectangularity} = \frac{\text{Area (Shape)}}{\text{Area (MBR)}}$$

در رابطه (۴) صورت کسر مساحت پوش محدب و مخرج کسر مساحت MBR می‌باشد.

o میزان دایره‌ای بودن شکل:[3] میزان دایره‌ای بودن شکل مطابق با رابطه (۵) محاسبه می‌شود [۲۶].

رابطه(۵)

$$\text{Circularity} = \frac{\text{Area}}{\text{Perimeter}^2}$$

o تعداد رئوس شکل: که بیانگر تعداد بازیکن قرار گرفته بر روی مرز شکل تیم می‌باشد.

• **توصیف‌کننده‌های توپولوژیکی:** این پارامترها که بر مبنای روابط توپولوژیکی و هپوشانی محدوده تیم با هندسه‌های توپ و زمین تعریف می‌شوند عبارتند از:

o مساحت محدوده مشترک با زون‌های زمین: مربیان برای مدیریت بهتر فضا، معمولا زمین را به زون‌های مشخصی تقسیم می‌کنند و از این زون‌بندی برای تعریف محدوده بازی بازیکنان و ترکیب‌بندی آن‌ها استفاده می‌کنند. شکل(۷) زون‌بندی‌های پیشنهادی توسط دو مربی حال حاضر فوتبال، لوئیس ون گال [۲۷] و پپ گواردیولا [۲۸] را نشان می‌دهد. اینکه در هر فریم زمانی، چه درصدی از مساحت تیم در هریک از این زون‌ها قرار می‌گیرد می‌تواند به عنوان پارامترهای خوشه‌بندی لحاظ شود.

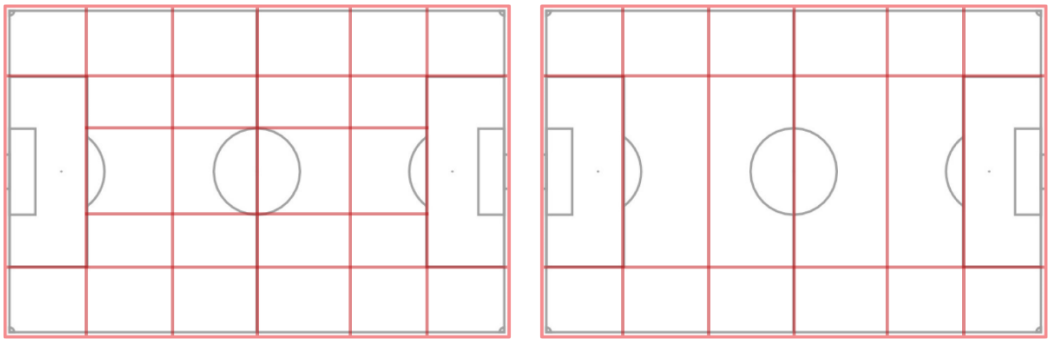
[1] *Elongation*

[2] *Rectangularity*

[3] *Circularity*

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]



(ب) (الف)

شکل۷: زون‌بندی زمین فوتبال به روش‌های مختلف، (الف) روش لوئیس ون گال ، (ب) روش پپ گواردیولا

- مجاورت با توپ: این که میزان نزدیکی تیم به توپ چقدر است، می‌تواند نقش مهمی در شکل تیم داشته باشد. از این رو مجاورت با توپ در قالب پارامتریک کمی فاصله تا توپ در محاسبات وارد می شود. این فاصله به صورت کمترین فاصله بین بازیکنان تیم تا توپ تعریف می‌شود.

- **توصیف‌کننده‌های تراکم و توزیع:** مساله مهمی که در ارتباط با هندسه پوش محدب برای تعریف شکل یک تیم وجود دارد، آن است که صرفا بر مبنای بازیکنان قرار گرفته در لبه تیم شکل می‌گیرد (و نه همه بازیکنان) و همچنین اطلاعاتی در ارتباط با تراکم و چگالی موقعیت بازیکنان در اختیار قرار نمی-دهد. برای رفع این چالش، فاصله بین مرکز هندسی شکل تیم و مرکز هندسی بازیکنان را به عنوان شاخصی برای میزان توزیع همگون از بازیکنان در محدوده تیم استفاده می‌شود.

۴-۲- گام دوم: اعمال روش خوشه‌بندی دو مرحله‌ای

پس از محاسبه پارامترهای مختلف توصیف‌کننده، فرآیند خوشه‌بندی انجام می‌شود. هدف از خوشه‌بندی یافتن ساختارهای مشابه برای شکل تیم در بازه‌های زمانی مختلف مسابقه می‌باشد. در این مقاله، از الگوریتم *K-Means* برای خوشه‌بندی استفاده می‌شود. این الگوریتم با وجود آنکه می‌تواند برای حجم زیادی از داده‌ها مورد استفاده قرارگیرد، از لحاظ سهولت پیاده-سازی، سرعت اجرا، بار محاسباتی و میزان حافظه مورد نیاز الگوریتم بهینه‌ای به شمار می آید [۲۹]. در [۳۰] الگوریتم‌های یادگیری ماشین برای تحلیل داده‌های موقعیت بازیکنان مورد بررسی قرار گرفته است. در پژوهش مذکور، در نهایت از بین روش‌های درخت تصمیم‌گیری، *SVM*، *Logistic regression* و شبکه عصبی، الگوریتم *K-Mean* به عنوان روش بهینه انتخاب شده است. در این روش ابتدا تعداد خوشه‌های مشخص می‌شود، سپس به ازای هر خوشه، یک مرکز اولیه انتخاب می‌شود. بر مبنای مجموعه‌ای از پارامترها، فاصله داده‌ها تا مرکز خوشه‌ها محاسبه می‌شود و بر مبنای نزدیکترین فاصله، تمامی داده به خوشه متناظر نزدیک اختصاص داده می‌شود. آنگاه میانگین وزن دار داده های هر خوشه به عنوان مرکز جدید خوشه انتخاب می‌شود. این فرآیند تا زمانی که تمامی داده‌های مشابه در یک خوشه قرار گیرند ادامه پیدا می‌کند [۲۹].

در این مقاله، از الگوریتم *K-Means* برای خوشه‌بندی اشکال تیم در دو مرحله استفاده می‌شود. در مرحله اول، خوشه‌بندی بر مبنای پارامترهای مکانی تیم صورت می‌گیرد که بیانگر وضعیت تیم در آن لحظه از مسابقه در یکی از حالت‌های دفاع، حمله، و انتقال بازی از دفاع به حمله یا از حمله به دفاع است. در مرحله دوم به ازای هر یک از خوشه‌های مکانی شناسایی شده در فاز نخست، بر مبنای پارامترهای هندسی، توپولوژیکی و

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]
[DOI: 10.61186/jgit.12.3.63]

تراکم که تمرکز بهتری بر شکل کلی تیم دارند، خوشه-بندی انجام می‌شود.

به منظور ارزیابی نتایج خوشه‌بندی، میزان شباهت اشکال مختلف تیم در هر خوشه با مرکز خوشه اندازه-گیری می‌شود. این اندازه‌گیری شباهت بر مبنای سه پارامتر مساحت، محیط، تعداد رئوس و مطابق با روابط(۶) الی (۸) انجام می‌شود [۳۱].

رابطه(۶)

$$Sim_{Area} = 1 - \frac{\left|Area_{Shape} - Area_{ClusterCenter}\right|}{\max(Area_{Shape}, Area_{ClusterCenter})}$$

که در رابطه (۶)، $Area_{Shape}$ مساحت شکل تیم و $Area_{ClusterCenter}$ مساحت مرکز خوشه است.

رابطه(۷)

$$Sim_{Perimeter} = 1 - \frac{\left|Perimeter_{Shape} - Perimeter_{ClusterCenter}\right|}{\max(Perimeter_{Shape}, Perimeter_{ClusterCenter})}$$

که در رابطه (۷)، $Perimeter_{Shape}$ محیط شکل تیم و $Perimeter_{ClusterCenter}$ محیط مرکز خوشه است.

رابطه(۸)

$$Sim_{Vertices} = 1 - \frac{\left|Vertices_{Shape} - Vertices_{ClusterCenter}\right|}{\max(Vertices_{Shape}, Vertices_{ClusterCenter})}$$

که در رابطه (۸)، $Vertices_{Shape}$ تعداد رئوس شکل تیم و $Vertices_{ClusterCenter}$ تعداد رئوس مرکز خوشه است.

با محاسبه میزان شباهت، می‌توان میزان تفاوت اشکال تیم در مقایسه با شکل متناظر با مرکز هر خوشه را با کمک رابطه (۹) محاسبه کرد:

رابطه(۹)

$$Diff_{shape\ i} = 1 - (w_a * Sim_{Area} + w_p * Sim_{Perimeter} + w_v * Sim_{Vertices})$$

که در رابطه (۹)، پارامترهای *Sim* بیان کننده میزان شباهت مبتنی بر مساحت، محیط و تعداد خوشه و پارامترهای w وزن متناظر با هریک می‌باشد. در این جا وزن‌ها برابر با هم در نظر گرفته می‌شوند.

۵- پیاده‌سازی

در این بخش به تشریح داده‌ها و نرم‌افزارهای مورد استفاده، ارائه و ارزیابی نتایج پیاده‌سازی، و تفسیر آن‌ها پرداخته می‌شود.

۵-۱- داده‌ها و نرم افزارهای مورد استفاده

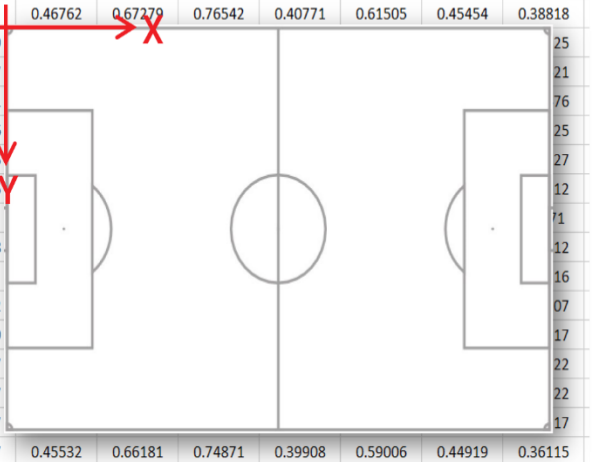
داده‌هایی که در این مقاله برای پیاده‌سازی روش پیشنهادی مورد استفاده قرار گرفته است، از منبع داده متریکا اسپرت دریافت شده است. این داده‌ها، شامل دو فایل موقعیت بازیکنان دو تیم و فایل رخدادهای مسابقه در فرمت *csv* می باشد. حجم هرکدام از فایل-های موقعیت بازیکنان در حدود ۲۸/۵ مگابایت و حجم فایل رخدادهای بازی در حدود ۱۳۰ کیلوبایت می‌باشد. طبق اطلاعات سایت مورد نظر، اطلاعات نام بازیکنان و اطلاعات تیم‌ها از فایل‌ها حذف شده و عملا داده‌ها ناشناس محسوب می‌شوند. ساختار داده‌های دریافتی مطابق با شکل (۸) می‌باشد. در این فرمت به ازای هر ۰/۴ ثانیه از بازی یک فریم زمانی در نظر گرفته شده است. از این رو در مجموع به ازای ۱۴۵۰۰۰ فریم از مسابقه، از موقعیت بازیکنان به همراه موقعیت توپ در یک سیستم مختصات محلی استفاده می‌شود.

داده‌های دریافتی، برای انجام تحلیل‌ها، به پایگاه داده اورکل *19c* وارد شده است. فرآیند پیش‌پردازش داده‌ها، ساخت پوش محدب از روی داده‌ها و همچنین محاسبه پارامترهای خوشه‌بندی با استفاده از قابلیت‌های مکانی اوراکل در این محیط انجام شده است. انجام فرآیند خوشه‌بندی به روش *K-Means* در محیط برنامه‌نویسی پایتون و با کتابخانه *Scikit Learn* صورت گرفته است. با اتمام فرآیند خوشه‌بندی، شماره خوشه متناظر با هر فریم زمانی در پایگاه داده بروزرسانی شده و برای نمایش نتایج خوشه‌بندی از نرم‌افزار *QGIS* استفاده شده است. در *QGIS* با اتصال به اوراکل، داده‌های هندسی بازیابی شده و نمایش داده شده است.

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]

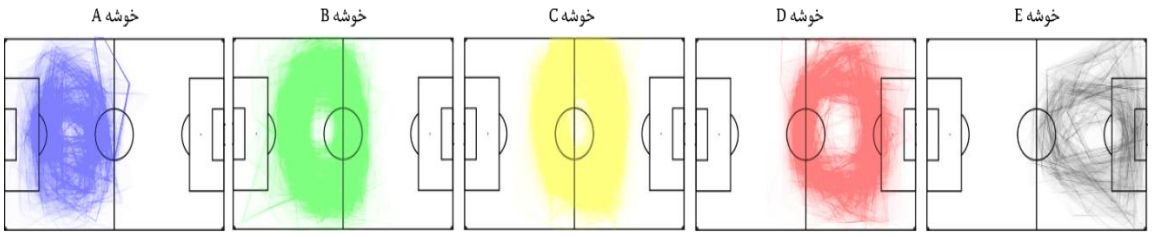
Period	Frame	Time	Player11_X	Player11_Y	Player1_X	Player1_Y	Player2_X	Player2_Y	Player3_X	Player3_Y	Player4_X	Player4_Y	Player5_X	Player5_Y
1	1	0.04	0.90509	0.47462	0.58393	0.20794	0.67658	0.4671	0.6731	0.76476	0.40783	0.61525	0.45472	0.38709
1	2	0.08	0.90494	0.47462	0.58393	0.20794	0.67658	0.4671	0.6731	0.76476	0.40783	0.61525	0.45472	0.38709
1	3	0.12	0.90434	0.47463	0.58393	0.20794	0.67658	0.4671	0.6731	0.76476	0.40783	0.61525	0.45472	0.38709
1	4	0.16	0.90377	0.47463	0.58351	0.20868	0.6764	0.46762	0.67279	0.76542	0.40771	0.61505	0.45454	0.38818
1	5	0.2	0.90324	0.47464	0.58291	0.21039	0.67599							25
1	6	0.24	0.90275	0.47464	0.58214	0.21123	0.67537							21
1	7	0.28	0.90227	0.47465	0.58142	0.21166	0.67471							76
1	8	0.32	0.90213	0.47374	0.58067	0.21182	0.67406							25
1	9	0.36	0.90204	0.4727	0.57998	0.21176	0.67335							27
1	10	0.4	0.90187	0.47188	0.57931	0.21151	0.67265							12
1	11	0.44	0.90163	0.47121	0.57873	0.21075	0.672							71
1	12	0.48	0.90134	0.47066	0.57816	0.21007	0.67138							12
1	13	0.52	0.90105	0.4702	0.5776	0.20929	0.6708							16
1	14	0.56	0.9007	0.46981	0.57701	0.20828	0.67032							07
1	15	0.6	0.90031	0.46947	0.57638	0.20744	0.66989							17
1	16	0.64	0.89987	0.46917	0.57574	0.20661	0.66947							22
1	17	0.68	0.89941	0.46892	0.57504	0.20556	0.66907							22
1	18	0.72	0.89893	0.46869	0.57428	0.20464	0.66867							17
1	19	0.76	0.89848	0.46849	0.57343	0.20364	0.66827	0.45532	0.66181	0.74871	0.39908	0.59006	0.44919	0.36115



شکل۸: فرمت خام داده‌های مورد استفاده در پیاده‌سازی

۵-۲- تحلیل داده‌ها

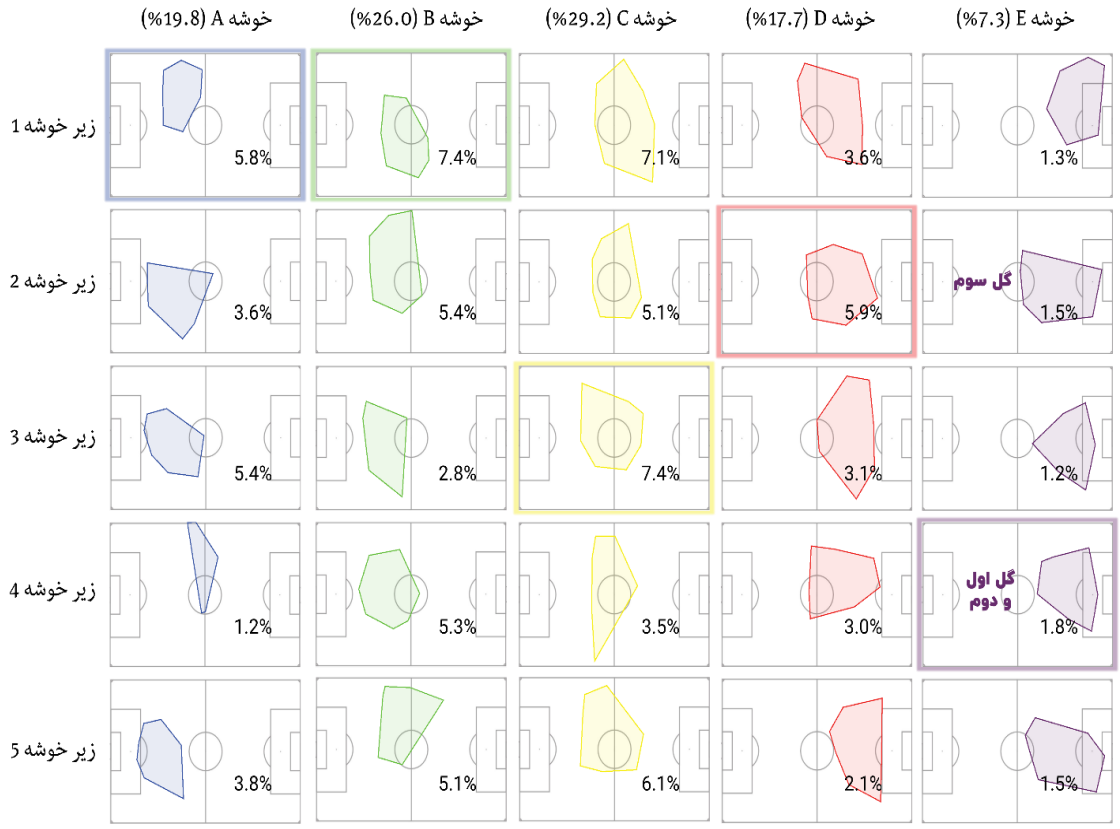
با انجام فرآیند آماده‌سازی و پیش‌پردازش بر روی این داده‌ها، پارامترهای توصیف‌کننده شکل در هر فریم زمانی محاسبه می‌گردد. سپس، برای آنکه بتوان خوشه-بندی معناداری از شکل تیم در وضعیت‌های مختلف داشت، خوشه‌بندی در دو مرحله انجام می‌شود. در مرحله اول، برای خوشه‌بندی به روش *K-Means*، تعداد خوشه‌ها برابر با ۵ و تعداد تکرار، ۲۰ در نظر گرفته می-شود. نتایج خوشه‌بندی انجام شده در این مرحله مطابق با شکل (۹) می‌باشد.



شکل۹: موقعیت تمامی شکل‌های تیم در خوشه‌های پنج گانه در مرحله اول پیاده‌سازی

در مرحله دوم نیز به ازای هر یک از خوشه‌های مکانی شناسایی شده در فاز نخست، فرآیند خوشه‌بندی با استفاده از روش *K-Means* و با در نظر گرفتن ۵ خوشه در ۲۰ تکرار انجام می‌شود. شکل (۱۰) خروجی‌های خوشه‌بندی شکل تیم در مرحله دوم را مشخص می‌کند. در این شکل، سهم زمانی هریک از خوشه‌ها و زیرخوشه‌ها مشخص شده است. در هر خوشه، زیرخوشه با بیشترین سهم زمانی متمایز شده است. همچنین زیرخوشه‌هایی که گل‌های مسابقه در آن زده شده است مشخص شده است.

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]
[DOI: 10.61186/jgit.12.3.63]



شکل۱۰: نتایج حاصل از خوشه‌بندی شکل تیم در مرحله دوم به همراه سهم زمانی هر خوشه در طول مدت مسابقه

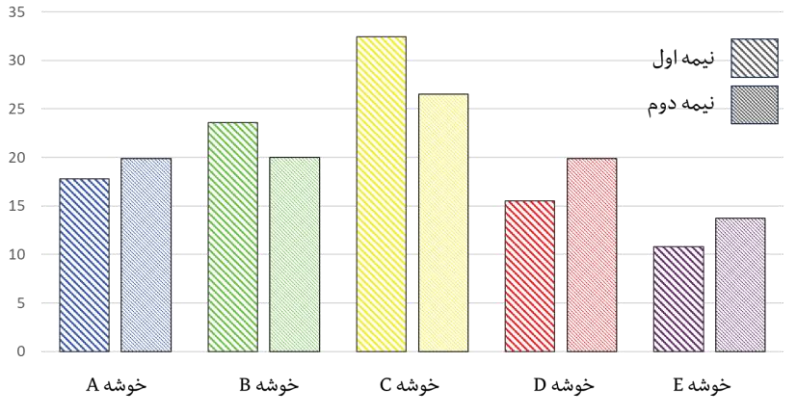
در بخش بعدی به تفسیر نتایج حاصل از خوشه‌بندی پرداخته شده است.

۵-۳- تفسیر نتایج

در مرحله اول از فرایند تحلیل، موقعیت تیم در یکی از ۵ منطقه افقی زمین خوشه‌بندی شده است. مطابق با شکل(۹)، خوشه‌های A، B، C، D و E به ترتیب بیانگر موقعیت تیم در محدوده دفاعی، نیمه زمین خودی، میانه زمین، نیمه زمین تیم حریف و محدوده دفاع تیم حریف می‌باشند. بر این اساس و بر مبنای سهم زمانی هریک از این خوشه‌ها می‌توان، از نتایج این خوشه‌بندی برای تحلیل استراتژی کلی تیم در بازی انجام شده بهره برد. مقایسه این سهم زمانی در بازه‌های زمانی مختلف می‌تواند وضعیت کلی تیم را مشخص کند. شکل(۱۱) نسبت فراوانی سهم زمانی هریک از خوشه‌های اصلی در طول هر یک از دو نیمه زمانی مسابقه را نشان می‌دهد.

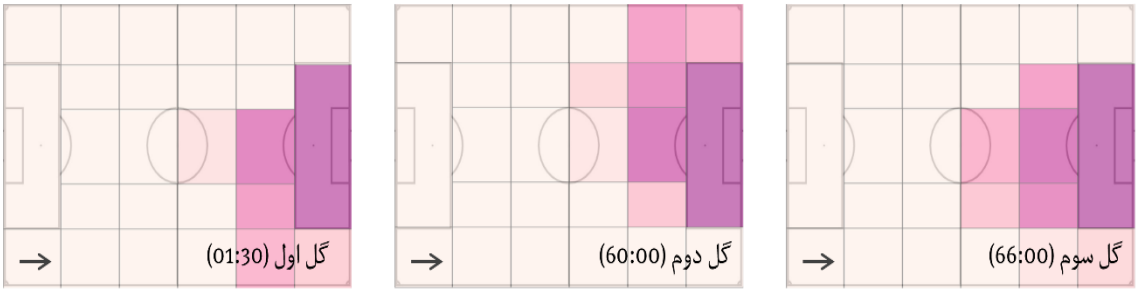
مطابق با شکل (۱۱)، سهم خوشه B و C در نیمه اول بیشتر از نیمه دوم است. این بدان معناست که تیم، بیشتر زمان بازی را در نیمه میانی زمین خود و میانه زمین حضور داشته است. طبق داده‌های رخدادهای بازی، تیم در دقیقه ۱:۳۰ بازی موفق به گل‌زنی شده است و از این رو می‌توان نتیجه گرفت که تلاش تیم در نیمه نخست بازی بیشتر بر حفظ نتیجه بوده است. این تلاش موفقیت‌آمیز بوده است و تیم حریف موفق به زدن گل تساوی نشده است. در نیمه دوم بازی، سهم خوشه A، D و E بیشتر از نیمه اول می‌باشد. پس حضور تیم در زمین تیم حریف بیشتر شده است و نتیجه این حضور، به ثمر رسیدن دو گل در دقایق ۶۰:۰۰ و ۶۶:۰۰ بازی بوده است. به موازات، فشار تیم حریف برای جبران گل‌های خورده شده بیشتر شده و همین امر سبب شده است که حضور تیم در محدوده دفاعی (خوشه A) نیز در این نیمه بیشتر شود.

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]
[DOI: 10.61186/jgit.12.3.63]



شکل۱۱: سهم زمانی خوشه‌های مرحله اول در هریک از دو نیمه مسابقه

شکل (۱۲) مجموع مساحت محدوده همپوشانی تیم با زون های تعریف شده در زمان به ثمر رسیدن سه گل بازی را نشان می‌دهد. با انطباق زمان‌های گل‌های بازی با خوشه‌بندی‌های تعریف شده، مشخص شده است که گل اول و دوم تیم در زمانی که تیم در زیرخوشه *E4* قرار داشته و گل سوم در زمانی که در زیرخوشه *E2* قرار داشته است، زده شده است. این دو زیرخوشه مطابق با شکل (۱۰) بیشترین سهم زمانی را در خوشه *E* به خود اختصاص داده است.

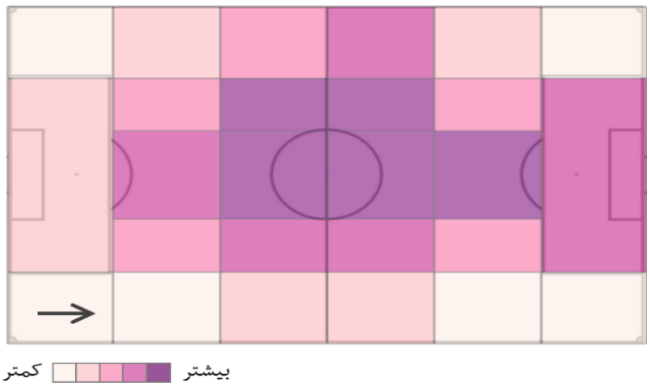


شکل۱۲: مجموع مساحت محدوده همپوشانی تیم با زون‌های تعریف شده در زمان به ثمر رسیدن سه گل بازی

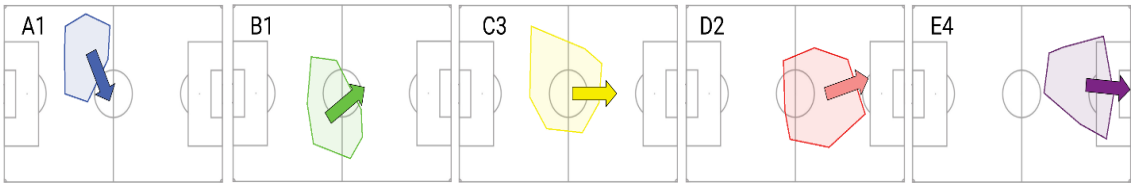
حال چنانچه مشابه آنچه در شکل (۱۲) نمایش داده شده است، سهم حضور تیم در زون‌های مختلف بازی در کل زمان بازی مشخص شود، شکل (۱۳) حاصل می‌شود. مطابق با این شکل، حضور تیم در محدوده میانی زمین بازی و همچنین محدوده مجاور محوطه جریمه تیم حریف بیشتر بوده و حضور کمتری در زون-های پایینی زمین (سمت راست تیم) داشته است. بر این معنای می‌توان حدس زد که بیشتر پیشروی‌های تیم به سمت دروازه تیم حریف از مسیر زون‌های بالایی زمین (سمت چپ تیم) انجام شده است.

برای بررسی دقیق‌تر این موضوع، زیرخوشه‌های با بیشترین سهم زمانی در هر خوشه در کنار یکدیگر قرار داده شده است. این توالی، مشابه آنچه در شکل (۱۴) آمده است، می‌تواند تا حدودی نمایی کلی مسیر حرکت رو به جلوی تیم در طول زمان مسابقه را ارائه دهد. این شکل می‌تواند موضوع پیشروی کلی تیم به سمت دروازه تیم حریف از مسیر زون‌های بالایی زمین را تایید کند.

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]
[DOI: 10.61186/jgit.12.3.63]

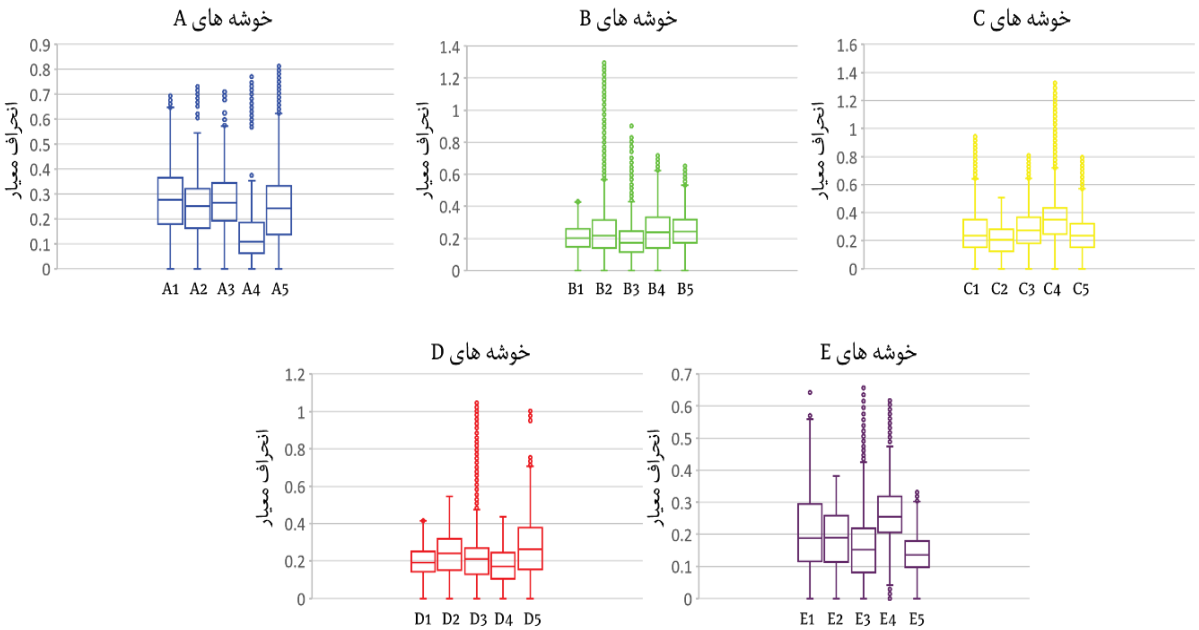


شکل۱۳: مجموع مساحت محدوده همپوشانی تیم با زون‌های تعریف شده در کل زمان بازی



شکل۱۴: مسیر حرکت کلی روی به جلوی تیم در طول مسابقه

در این مقاله پس از انجام خوشـه‌بنـدی، میـزان تفـاوت اشکال تیم در طول زمان مسابقه نسبت به مرکز خوشه متنـاظر، مطـابق بـا روابـط (۶-۹) محاسـبه مـی‌شـود. میانگین و انحراف معیار این میزان تفاوت در شکل(۱۵) نمایش داده شده است.



شکل۱۵: میزان تفاوت اشکال تیم با مراکز خوشه‌ها

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]
[DOI: 10.61186/jgit.12.3.63]

مطابق با این شکل متوسط میزان انحراف معیار در زیرخوشه‌های D و E در مقایسه با سه خوشه دیگر کمتر است. هرچه انحراف معیار کمتر باشد، تنوع شکل-های موجود در آن خوشه بر مبنای مساحت، محیط و تعداد رئوس کمتر خواهد بود. در تفسیر این موضوع می‌توان گفت که تیم در این زمین تیم حریف، الگوهای تقریبا یکسانی را برای مدیریت فضای بازی و پیشروی مورد استفاده قرار داده است. قرار گرفتن شکل تیم در زمان دو گل از سه گل بازی در یک زیر خوشه هم می-تواند این موضوع را تایید کند. اما در میانه زمین (زیر خوشه‌های C) و زمین خودی (زیر خوشه‌های A و B) نوسانات تغییر شکل تیم بیشتر بوده است و تیم استراتژی‌های متنوع‌تری را برای دفع حمله تیم حریف در پیش گرفته است.

در این جا لازم به یادآوری است که خوشه‌بندی شکل تیم، موضوعی است که در پژوهش‌های دیگر نیز به آن پرداخته شده است. یکی از مشابه ترین خروجی‌هایی که از خوشه‌بندی شکل تیم به دست آمده است، نتایجی است که مربوط به پژوهش [۹] می‌باشد. در پژوهش مذکور خوشه‌بندی شکل تیم صرفا برمبنای خصوصیات هندسی و حتی مستقل از مکان، دوران و مساحت تیم انجام شده است. این امر در حالی که در مقاله پیش رو، در نظر گرفتن پارامترهای بیشتر و همچنین دومرحله‌ای نمودن فرآیند خوشه‌بندی، امکان ارائه تفسیرهای معنادارتری از نتایج از جمله بررسی رفتار کلی تیم در محدوده‌های مختلف زمین را فراهم نموده است.

۶- نتیجه‌گیری

در این مقاله با تمرکز بر رو شکل کلی تیم به جای تحلیل بازی تک تک بازیکنان سعی شده است یک بینش کلی از رفتار تجمعی تیم در طول یک مسابقه ارائه شود. این تحلیل بر مبنای داده‌های زمانی-مکانی بازیکنان انجام گرفته است. لازمه این کار در نظر گرفتن یک شکل هندسی کلی برای تیم در هر لحظه می‌باشد. هندسه‌ای که برای این امر در نظر گرفته شده است، پوش محدبی است بر مبنای موقعیت بازیکنان تیم (به جز دروازه بان) ساخته شده است. سپس برای توصیف شکل تیم در هر زمان، پارامترها و شاخص‌هایی محاسبه و استخراج شده است. این پارامترها شامل پارامترهای مکانی، هندسی، توپولوژیکی، و توزیع و تراکم بازیکنان می‌باشد. بر اساس این پارامترها، فرآیند خوشه‌بندی شکل تیم به روش *K-Means* در دو مرحله انجام شده است. در مرحله اول پنج خوشه اصلی شناسایی می‌شود که موقعیت تیم را در محدوده‌های مختلف زمین مشخص می‌کند. آنگاه در مرحله دوم به ازای هر خوشه، پنج زیر خوشه تعریف می‌شود که بر مبنای آن می‌توان الگوهای هندسی مشابه برای شکل تیم در طول زمان مسابقه را تعیین کرد. بررسی نتایج این خوشه‌بندی‌ها، می‌تواند اطلاعات ارزشمندی را برای شناسایی و تجزیه و تحلیل تیم‌ها در اختیار مربیان قرار دهد. کشف الگوهای تکرار شونده از شکل تیم به همراه سهم زمانی آن‌ها، شناسایی مناطق با بیشترین و کمترین میزان تسلط تیم بر آن‌ها، استخراج الگو کلی حرکت روبه جلوی تیم و میزان نوسانات شکل تیم در مناطق مختلف زمین از جمله نتایجی است که از تفسیر خوشه‌بندی‌ها به دست می‌آیند.

با وجود دست‌یابی به این نتایج، این مقاله بیشتر بر تحلیل داده‌های مکانی موقعیت بازیکنان و مشخصات هندسی شکل تیم متمرکز شده است و مواردی همچون بررسی تاثیر رخدادهای مختلف مسابقه بر روند بازی و یا نحوه تعامل تیم با تیم حریف کمتر مورد توجه قرار گرفته است. افزون بر این با وجود مزایای پوش محدب برای تعریف شکل تیم، موقعیت بازیکنان قرار گرفته در داخل مرز پوش محدب در تحلیل‌ها نادیده گرفته شده است. بر این اساس پیشنهاد می‌شود که در تحقیقات آتی ضمن بازنگری در تعریف شکل تیم، رخدادهای بازی و رفتار متقابل تیم حریف هم مد نظر قرار گیرند. همچین با انجام تحلیل‌ها بر روی داده‌های مسابقات بیشتری از یک تیم، می‌توان الگوهای بیشتر و دقیق-تری از تیم را شناسایی کرد.

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]

مراجع

[1]. F.R. Goes, L.A. Meerhoff, M.J.O. Bueno, D.M. Rodrigues, F.A. Moura, M.S. Brink, M.T. Elferink-Gemser, A.J. Knobbe, S.A. Cunha, R.S. Torres, and K.A.P.M. Lemmink, "Unlocking the Potential of Big Data to Support Tactical Performance Analysis in Professional Soccer: A Systematic Review", European Journal of Sport Science. Eur J Sport Sci, Vol. 21(4), pp. 481-496, 2020, DOI: 10.1080/17461391.2020.1747552.

[2]. J. Gudmundsson, and T. Wolle, "Football Analysis Using Spatio-Temporal Tools", Computers, Environment and Urban Systems, Vol. 47, pp. 16-27, 2014, DOI: https://doi.org/10.1016/j.compenvurbsys.2013.09.004

[3]. J. Gudmundsson, and M. Horton, "Spatio-Temporal Analysis of Team Sports", ACM Computing Surveys, Vol. 50(2). pp 1–34, 2017, DOI:10.1145/3054132

[4]. Y. Li, R. Ma, B. Gonçalves, B. Gong, Y. Cui, and Y. Shen, "Data-driven team ranking and match performance analysis in Chinese Football Super League", Journal of Chaos, Solitons & Fractals, Vol 141 (110330), 2020, DOI: https://doi.org/10.1016/j.chaos.2020.110330

[5] D. Araújo, P. Passos, P. Esteves, R. Duarte, J. Lopes, R. Hristovski, and K. Davids, "The Micro-Macro Link in Understanding Sport Tactical Behaviours: Integrating Information and Action at Different Levels of System Analysis in Sport", Movement & Sport Sciences – Science & Motricité, pp. 53–63, 2015, DOI:10.1051/sm/2015028

[6]. V. Kaldaras, Y. Michailidis, I. Gissis, and T.I. Metaxas, "The Running Performance of Amateur Football Players in Matches with a 1-4-3-3 Formation in Relation to Their Playing Position and the 15-min Time Periods", Appl. Sci., Vol. 14 (7036), 2024, DOI: https://doi.org/10.3390/app14167036

[7]. R. Izzo, S. Franco, and C. Hosseini Varde'I, "Analysis of Speed Thresholds in Youth Amateur Football Players Divided by Roles Using GPS Technologies", Journal of Sports Sciences, Vol. 6, 2018, DOI: 10.17265/2332-7839/2018.04.004

[8]. M. Mitrotasios, I. Ispyrlidis, N. Mantzouranis, G. Paraskeyopoulos, "GPS-based performance analysis in amateur football: Match Evaluation Relative to Players' Tactical Roles", Journal of Physical Education and Sport, Vol. 24(5), pp. 1275-1280, 2024, DOI:10.7752/jpes.2024.05145

[9]. M.J.O. Bueno, M. Silva, S.A. Cunha, R.S. Torres, F.A. Moura, "Multiscale Fractal Dimension Applied to Tactical Analysis in Football: A Novel Approach to Evaluate the Shapes of Team Organization on The Pitch", PLoS ONE, Vol. 16(9), 2021, https://doi.org/10.1371/journal.pone.0256771

[10]. G. Zhang, "The Analysis of Team Tactical Behaviour in Football Using GNSS Positional Data", Master Thesis, Liverpool John Moores University, 2022, DOI:10.24377/LJMU.t.00019091. https://researchonline.ljmu.ac.uk/id/eprint/19091/1/2022guangzezhangmphil.pdf

[11]. L. Shaw, and M. Glickman, "Dynamic Analysis of Team Strategy in Professional Football", BARÇA Sports Analytics Summit, 2020, https://static.capabiliaserver.com/frontend/clients/barca/wp_prod/wp-content/uploads/2020/01/56ce723e-barca-conference-paper-laurie-shaw.pdf

[12]. J. Whitmore, and T. Seidl, "Shape Analysis: Automatically Detecting Formations", 2021, https://theanalyst.com/na/2021/03/shape-analysis-automatically-detecting-formations/

[13]. B. Goncalves, D. Coutinho, J. Exel, B. Travassos, C. Lago, and J. Sampaio,

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]

"Extracting Spatial-Temporal Features That Describe a Team Match Demands When Considering the Effects of The Quality of Opposition in Elite Football", PLoS ONE, Vol. 14(8), 2019, https://doi.org/10.1371/journal.pone.0221368

[14]. T. Narizuka, and Y. Yamazaki, "Clustering Algorithm for Formations in Football Games", Scientific Reports, Vol. 9(13172), 2019, https://doi.org/10.1038/s41598-019-48623-1

[15]. A. Bialkowski, P. Lucey, P. Carr, Y. Yue, S. Sridharan, and I. Matthews, "Large-Scale Analysis of Soccer Matches Using Spatiotemporal Tracking Data", presented at IEEE International Conference on Data Mining, Shenzhen, China, pp.725-730, 2014, DOI:10.1109/ICDM.2014.133

[16]. N. Liu, "Geovisualisation of Football Players Movement", Diploma Thesis, Palacký University Olomouc, Faculty of Science, Department of Geoinformatics,2022, https://theses.cz/id/mbeof5/48261609

[17]. J.V. Haaren, P. Robberechts, T. Decroos, L. Bransen, and J. Davis, "Analyzing Performance and Playing Style Using Ball Event Data", 2019, https://www.janvanhaaren.be/assets/papers/bih-2019-event-data.pdf

[18]. StatsBomb Open Data: https://github.com/statsbomb/open-data

[19]. Football Events from Kaggle: https://www.kaggle.com/datasets/secareanualin/football-events

[20]. Soccer match event dataset in Wyscout: https://doi.org/10.6084/m9.figshare.c.4415000.v5, https://doi.org/10.6084/m9.figshare.9711164

[21]. Metrica Sports Sample Data: https://github.com/metrica-sports/sample-data

[22]. FBref Website: https://fbref.com/en

[23]. A. Fujimura, and K. Sugihara, "Geometric Analysis and Quantitative Evaluation of Sport Teamwork", Systems and Computers in Japan, Vol. 36(6), 2005, DOI:10.1002/scj.20254

[24]. H. Fan, Z. Zhao, and L. Wenwen, "Towards Measuring Shape Similarity of Polygons Based on Multiscale Features and Grid Context Descriptors", ISPRS Int. J. Geo-Inf., Vol. 10(279), 2021, https://doi.org/10.3390/ijgi10050279

[25]. M.A. Wirth, "Shape Analysis & Measurement", University of Guelph, Computing and Information Science, Image Processing Group, 2004, http://www.cyto.purdue.edu/cdroms/micro2/content/education/wirth10.pdf

[26]. F. Park, "Shape Descriptor/Feature Extraction Techniques", UCI iCAMP, 2011, https://www.researchgate.net/profile/James-Peters-3/post/How-to-extract-features-for-classification-of-geometry-design/attachment/59d62643c49f478072e9ade8/AS%3A272177448849413%401441903568230/download/shape-descriptors.pdf

[27]. Six-by-Three Grid: https://spielverlagerung.com/glossary/pitch-zones/six-by-three/

[28]. Positional play: football tactics explained: https://www.coachesvoice.com/cv/positional-play-football-tactics-explained-guardiola-cruyff-manchester-city/

[29] L. Morissette, and S. Chartier, "The K-Means Clustering Technique: General Considerations and Implementation in Mathematica", Tutorials in Quantitative Methods for Psychology. Vol. 9(1), p. 15-24. 2013, DOI: 10.20982/tqmp.09.1.p015

[30]. R. Gasparini, and A. Álvaro, "Positional Analysis of Brazilian Soccer Players Using GPS Data", Revista Brasileira de Computação Aplicada, Vol. 12(3), pp. 16–32, 2020, DOI: 10.5335/rbca.v12i3.10234

[31]. R. Duraciová, "An Aggregated Shape Similarity Index: A Case Study of

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]

Comparing the Footprints of OpenStreetMap and INSPIRE Buildings", ISPRS Int, J.Geo-Inf, Vol. 12 (495), 2023, DOI:.org/10.3390/ijgi12120495

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]

K. N. TOOSI UNIVERSITY OF TECHNOLOGY
FACULTY OF GEODESY AND GEOMATICS ENGINEERING

Journal of Geospatial Information Technology
Vol.12, No.3, Autumn 2024

Research Paper

Shape Clustering Based on Spatio-Temporal Data for Analyzing the Collective Behavior of a Football Team Players

Ali Zare Zardiny *, Zahra Bahramian

Assistant professor in School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran

Abstract

In recent years, the analysis of football data with the aim of investigating the behavior of players has received a lot of attention. An important part of these data has a spatio-temporal nature, and this makes the importance of spatio-temporal analyzes more prominent in the football industry. The purpose of this research is to analyze the collective behavior of the players at the macro level. For this purpose, in the first step in each time frame, the team characteristics are extracted based on a set of spatial, geometric, topological and distribution parameters. Then, these parameters are the basis of the team shape clustering. This clustering is performed in two phases. In the first phase, the main clusters are obtained based on the spatial parameters and the defensive or offensive status of the team is determined. In the second phase, for each of the main clusters, based on some other descriptors, new sub-clusters are defined. In this research, the data of a football match was used and five main clusters, as well as five sub-clusters for each main cluster, were identified. In the evaluation process, the difference between the shape of the team and the center of the corresponding cluster has been measured. The standard deviation of this difference in the main clusters varies between 0.19 and 0.27. Based on this change in the standard deviation, the fluctuations in the team's shape in different areas of the pitch will be determined. And based on the clustering of the team's overall shape at different times of the match, the time contribution of the clusters and the degree of the team dominance on the pitch, the defensive or offensive situation of the team and also, the overall flow of the team's movement will be determined. Taking the spatial, topological and density parameters along with the geometrical ones into consideration, performing clustering in two stages and not needing to transfer the shape to the raster space (and as a result not needing to use the image processing techniques) are the most important distinguishing traits of the proposed method compared to the previous researches.

Key words: *Analysis of collective behavior of football players, spatio-temporal data, convex hull, two-phase clustering, K-Means algorithm.*

Correspondence Address*: GIS Group, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran.*
Tel*: +98 912 4527859.*
Email: *zare_zardiny@ut.ac.ir*

[Downloaded from jgit.kntu.ac.ir on 2026-11-11]

[DOI: 10.61186/jgit.12.3.63]