The identification of subgroups using clustering analysis during successful possessions in football

SHUHEI, ISHIWATARI[1]. & SIGRID, B.H. OLTHOF[1].

[1]School of Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, UK

Football is characterised by the complex and dynamic nature where each player is moving with respect to team members, opponents, and the ball. Subgroups are formed where players work together to form a collective tactical behaviour of a team. Subgroup-level analysis using clustering techniques may be promising to differentiate successful and unsuccessful attacks, and to understand patterns in synchrony during successful attacks (Goes *et al.,* 2020, Journal of Sports Sciences, 39, 1-10). To date, subgroup analyses only included clusters with a fixed number of players, where dynamic subgroups could meet the dynamic nature of football. Therefore, the aim of the research was to explore the contribution of players to subgroups during successful attacks in football using a clustering technique. With institutional ethics approval, 22 elite-level under-19 male football players (mean ± SD: age 18.1 ± 1.0 years, height 178.4 ± 8.5 cm, weight 72.8 ± 8.1 kg) playing at a professional academy in the Netherlands were recruited to complete a 90-min football game. Positional data were collected using a Local Positioning System of 10hz to provide X and Y coordinates for each player. First, K-nearest neighbours search was conducted for parameter tuning of the clustering algorithm on the positional data (Rahmah and Sitanggang, 2016, *IOP conference series,* 31*,* 12012). Second, density-based spatial clustering of applications with noise (DBSCAN) was used to dynamically identify subgroups over the duration of the game. Successful attacks were defined as ball possessions resulting in shots or goals. From a total of 117 ball possessions, 12 attacks were classified as successful (*n*=10 attacks from open-play; *n*=2 from set-pieces). The remaining 105 attacks were unsuccessful. Successful attacks resulted in the mean number of 1.7 ± 0.4 clusters. On 7 out of 12 attacks, the number of clusters increased from the beginning of the possessions towards the end. On 7 out of 10 occasions excluding set-pieces, the striker was regarded as an outlier. Similarly, on 9 out of 10 occasions, wingers were classified as outliers. On all the occasions, the defensive midfielder was a part of clusters. The current results suggested that the unsupervised clustering algorithm successfully identified dynamic subgroups and outliers during a football match. Subgroups were evolving during ball possession where different players contributed to a successful outcome. Strikers and wingers were classified as outliers in the majority of attacks, highlighting their unpredictable attacking role where explosiveness and creativity is required. The defensive midfielder played a vital role in all attacks, reflecting their pivot role to link attackers and defenders. This further reinforces the tactical understanding of football.