

# Final Project Proposal

CS4195 - Group 11

## Dataset

Temporal network data of players' activity in soccer matches.

Nodes: players

Links: Interaction between players (mostly passes)

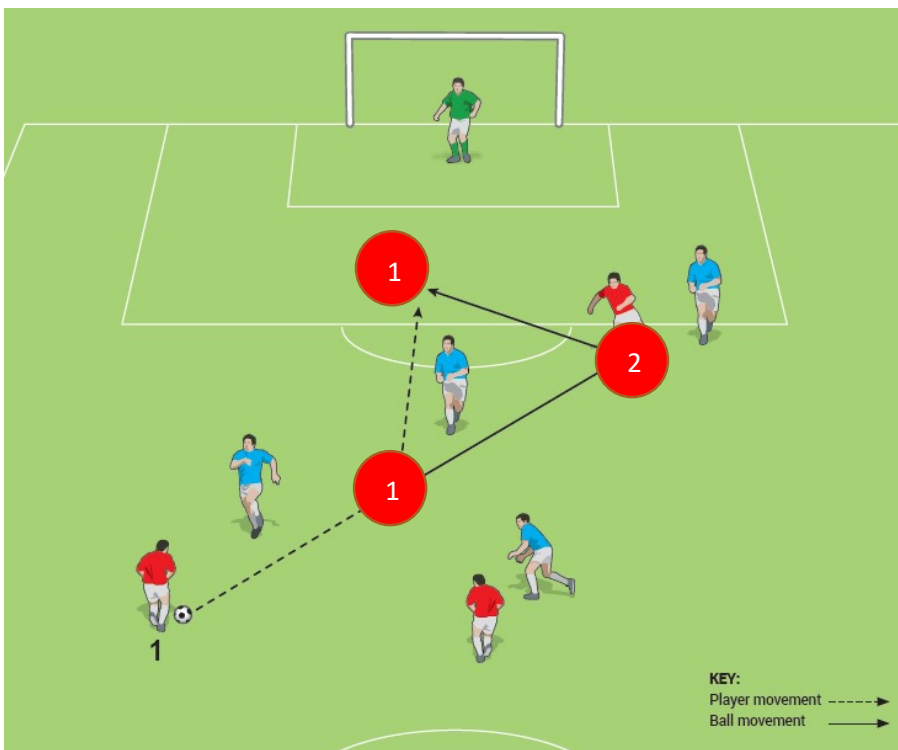
Data scale: entire season matches of seven competitions (containing event logs in every single match)

Reference: Pappalardo, Luca; Massucco, Emanuele (2019): Soccer match event dataset. figshare. Collection.  
<https://doi.org/10.6084/m9.figshare.c.4415000.v5>

Data: entire season matches of seven competitions

## What do we want to obtain from such a network?

1. Analyze players' performance and influence in soccer matches using graph metrics (network characterization)
2. Evaluate team performance and team play styles using graph metrics (network characterization)
3. Topological metrics, spectral metrics, centrality metrics, can they, and how can they characterize a soccer match? Can they distinguish strong and weak teams?
4. Can we use motifs to crack the coach's offensive playbook in temporal networks? (Motifs in temporal network can represent certain tactics, simple examples: one-two attacking tactic, see picture below)
5. Can we find the key to goals/winning in soccer games and be able to predict the result of a soccer match? (Perhaps by cross-correlation of metrics and game results/seasonal rankings)
6. (Optional) Can we identify players that are important to winning in reality but are not great from a traditional statistical perspective? How exact do they influence the match and what's missing in the regular statistics? (Key player to the team but probably have few goals/assists/passes/etc.)



A one-two is the classic pass, run and receive back move that bypasses the defender

