

Learning the Value of Teamwork to Form Efficient Teams

Ryan Beal,¹ Narayan Changder,² Timothy J. Norman,¹ Sarvapali D. Ramchurn¹

¹School of Electronics and Computer Science, University of Southampton,
Southampton, SO17 1BJ, United Kingdom
{ryan.beal, t.j.norman, sdr1}@soton.ac.uk

²The National Institute of Technology, Durgapur, West Bengal 713209, India
narayan.changder@gmail.com

Abstract

In this paper we describe a novel approach to team formation based on the value of inter-agent interactions. Specifically, we propose a model of teamwork that considers outcomes from chains of interactions between agents. Based on our model, we devise a number of network metrics to capture the contribution of interactions between agents. This is then used to learn the value of teamwork from historical team performance data. We apply our model to predict team performance and validate our approach using real-world team performance data from the 2018 FIFA World Cup. Our model is shown to better predict the real-world performance of teams by up to 46% compared to models that ignore inter-agent interactions.

1 Introduction

Team Formation (TF) is a fundamental concept that underpins many multi-agent systems where heterogeneous agents with individual properties (e.g., roles, capabilities, costs) come together to undertake tasks. TF involves the evaluation of different sets of agents in order to determine how well they will, individually or collectively, perform their tasks. By so doing, it is then possible to pick sets of agents that form the most effective teams. For example, teams of emergency responders are formed based on individual agent's abilities to navigate a difficult environment or address threats (Chalkiadakis and Boutilier 2012). Similarly, in ride-sharing settings, groups of riders can be efficiently formed to minimise travel time and costs (Bistaffa et al. 2017b).¹

Existing TF algorithms, such as (Fitzpatrick and Askin 2005), create effective human teams in the workplace using a mathematical programming formulation and a heuristic solution. Also, in (Scerri et al. 2005) a task allocation algorithm is discussed for extreme teams in disaster response. The models and algorithms in these papers have shown to be successful in their domains. However, these models typically ignore the fact that sets of agents interact in very specific ways. For example, agents in a team may transfer partly finished products to each other along a production line or a

firefighter may secure a building first before sending in a medic. In this paper, we consider how such *directed interactions* between agents can be valued and considered in the prediction of team performance.

Against this background, we propose a novel approach to forming teams using patterns that appear in a network of interactions between agents. We then validate and evaluate our approach by applying our models and algorithms to a real-world team formation problem presented by the domain of football (soccer). We show that our teamwork-focused model outperforms other player-focused approaches at predicting the teams that would be chosen by human-expert managers across 64 games from the 2018 FIFA World Cup. We also show that our model is better at predicting the performance of teams from real-world data. Thus, this paper advances the state of the art in the following ways:

1. We propose a novel approach to team formation based on the value of inter-agent interactions. Specifically, we propose a model of teamwork that considers the outcomes of the chains of such interactions.
2. Based on our model, we propose a number of network metrics to capture the contributions of individuals and sets of agents.
3. We show how the value of teamwork can be learnt from data and then applied to the prediction of team performance.

When taken together, our results establish the first benchmarks for team formation based on the learnt value of teamwork. Furthermore, our work opens up a new area of research into the use of teamwork-based models to understand how human teams work.

The rest of this paper is organised as follows. In Section 2 we review the literature, while Section 3 defines the network model and the optimisation problem. Section 4 discusses the application of our model to football to test and validate and Section 5 provides the detail of methods that we use to apply the model to the problems posed by football. We perform a number of Experiments on our model in Section 6 and discuss our findings in Section 7. Finally, Section 8 concludes.

¹Note that TF is different from coalition formation in terms of its focus on inter-agent interactions and non-selfish agents.

2 Related Work

There are examples in multi-agent systems literature which form teams using analysis of agents within a network. Gaston and DesJardins (2005), propose several strategies for agent-organised networks and evaluate their effectiveness for increasing organisational performance. The same authors also present an agent-based computational model of team formation, and analyse the theoretical performance of team formation in two simple classes of networks - ring and star topologies (Gaston and DesJardins 2008). Recently, (Bistaffa et al. 2017b) proposed a cooperative game-theoretic approach to deal with the problem of social ridesharing. They first formed a social network representation of a set of commuters, then proposed a model to form the coalition and arrange one-time rides at short notice. The authors model their problem as a Graph-Constrained Coalition Formation (GCCF) (Bistaffa et al. 2017a). Their model is based on two principles, first they solve the optimisation problem for making coalitions while minimising the cost of the overall system. The set of feasible coalitions in their model is restricted by a graph (i.e., the social network). Secondly, they address the payment allocation aspect of ridesharing.

Farinelli et al. (2004) survey the work that has been completed on the ability of the robots cooperation and coordination. There has also been work exploring how “ties” in social networks affect teamwork and the roles within teams, shown in (Balkundi and Kilduff 2006; Zhong et al. 2012). However, ties only address the closeness of agents within a social network, whereas we focus on how the frequency and success of interactions within a team impact the performance and outcomes.

To our knowledge, none of the discussed approaches have looked at directed interactions between team members and how chains of interactions lead to different team outcome events. More importantly, these approaches have only been validated on synthetic data. Instead, our work is validated on granular data about team performance in real-world games involving teams of humans.

3 The Team Formation Model

Our TF model is based on our observations of many real-world team-based systems, such as football teams or teams of emergency responders, as follows:

- Many teams operate through directed (one-to-one, one-to-many, or many-to-many) interactions. For example, in a football team, a player would pass a ball to another. However, team members will not always interact equally with every other team member. In this paper, as a first step, we will focus on one-to-one interactions. Indeed, we show that such a setup gives rise to complex interactions that pose difficult computational challenges.²
- Team members may have different roles and abilities to perform tasks. For example, emergency response teams will have members with different skill sets, equipment,

²We will consider more complex forms of interactions in future work.

and training. In a football team, each player will have a position on the pitch and specific abilities.

- Team actions can have multiple consequences. In the simplest case, they may have binary outcomes (succeed in achieving a mission or failing to do so). In many cases, however, team success is more nuanced (e.g. the achievement of a sub-ideal goal).
- Team formation typically involves picking a subset of agents that work well together, using some metric of efficiency. For example, emergency responders will choose a subset of available partners that are most fitting to the task or have the right skills. Similarly, a football manager will pick the best team (measured by their likelihood to win a match) of 11 players out of a squad of 23.

As can be seen, choosing the optimal team can be a difficult task given the complexity of the roles and relationships among team members as well as the environment they evolve in. In what follows, we formally define the key constructs of our teamwork model and devise multiple network-based metrics over which the value of teamwork can be learnt and used to predict the performance of different teams.

3.1 Basic Definitions

We define the set of n agents as $\mathcal{A} = \{a_1, a_2, \dots, a_n\}$. Agents can interact with each other to achieve some overarching goals. We consider such interactions to be directed (e.g., a UAV allocating a task to another UAV or a player passing a ball to another player in a football team). We define the set of interactions I as a set of ordered pairs (a_y, a_z) , where a_y initiates the interaction towards a_z . The same pair of agents may engage in such interactions multiple times, and hence I is effectively a multi-set of interactions.

Agents’ interactions are constrained by a directed graph representing potential roles and relationships between the agents. In the general case, all nodes in the graph will have at least two directed edges (one outgoing, and one incoming) between them and another agent. We denote this graph as $G = (\mathcal{A}, \mathcal{I}, w(\cdot))$ with agents $|\mathcal{A}|$ as vertices and edges \mathcal{I} representing relationships between pairs of agents. The weight on each edge is the number of directed interactions from the set I between pairs of agents. For example, for each edge $i \in \mathcal{I}$, the weight of i , $w(i)$ is defined as the number of times the edge $i = (a_x, a_y)$ appears in the multi-set I (the multiplicity of i in I). Formally, $w(i) = \sum_{s \in I} \mathbf{1}_{\{i\}}(s)$. This is an iteration over all elements of I and if $i = s$ a value of 1 is added (equivalently, if $i \in I$) and 0 otherwise. We build on these definitions to model how interactions between agents result in specific events.

3.2 Modelling Chains of Interactions

In many situations, agents will interact sequentially with each other (i.e., agent a_x interacts with a_y who in turn interacts with a_z). In this paper, we only consider the cases where an agent interacts with one other agent at a time.³ To

³We will consider cases where this assumption does not apply in future work. For example a ship may deploy many sub-ships who have all been allocated the same task to complete.

this end, we define a walk in the graph G as a sequence of interactions over the edges of the graph. Formally, a walk \mathcal{P} of length l in the graph G from vertex a_y to vertex a_z is a sequence $[a_x, a_y, \dots, a_z]$ where $l = |[a_x, a_y, \dots, a_z]|$. An example of a walk \mathcal{P} is described for football in Section 4.2 and this is shown in Figure 2. Another example of a walk in a real-world application would be the movement of a data-packet through a mesh network where the packet moves from router to router until it reaches the destination.

A walk leads to an event of a specific type. For example, a data packet being used to complete a file download, or a football player scoring a goal at the end of a sequence of passes. There may be many different event types. Formally, the set of possible events \mathcal{E} is defined as $\mathcal{E} = \{e_1, e_2, \dots, e_k\}$ where e_k is the event and k is the number of possible events from the walk.

Each of the possible events e_k may have a different impact on the environment, therefore affecting the overall performance of the team. Thus, for each $e \in \mathcal{E}$, the value function $\alpha : \mathcal{E} \rightarrow [0, 1]$ determines an associated value. For example, in a game of football if the event e_k is a ‘‘goal’’ event, this has a bigger impact on the overall outcome of the game and team performance in comparison to if e_k is a ‘‘loss of possession’’ event. Another example could be walks leading to a ‘‘person saved’’ event, in an emergency response setting, having greater impact than walks not leading to such events.

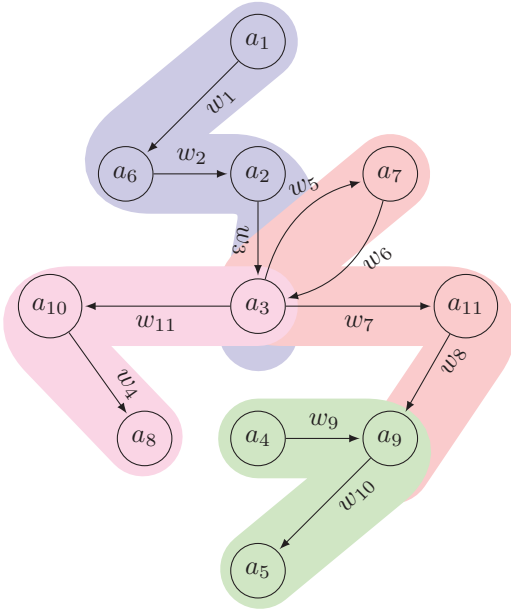


Figure 1: An example of 4 walks through a sample graph of 11 agents for an event e_k . The directed edge between two vertex represents the interaction between them and each highlighted colour represents a walk.

Note that each walk originates from one agent and involves chains of directed interactions between pairs of agents, resulting in an event. Hence, we next propose methods to extract the contribution of each agent as well as sets of agents to these individual events.

3.3 Extracting the Value of Teamwork

Walks \mathcal{P} and associated events \mathcal{E} can be used to infer the value of agents and sub-sets of agents within the team. We propose three metrics to value the contribution of individual agents and sets of agents as follows:

- **Centrality:** $v_{cent} : \mathcal{A} \rightarrow \mathcal{Z}$ refers to the sum of the weight of edges incident (incoming and outgoing) to a_i . This measures the influence of agent a_i in the network.
- **Distance from event:** $v_{dist} : \mathcal{A} \rightarrow \mathcal{Z}$ defines the average distance of agent a_i for each event. This represents the influence of agent a_i on an event.
- **Walk frequency:** $v_{freq} : 2^{\mathcal{A}} \rightarrow \mathcal{Z}$ refers to the number of times an agent a_i or a subset of agents appears in all walks. This represents the influence of an agent in the team.

It is important to note that these metrics attempt to summarise team performance in different ways, each with a different degree of information loss. Using centrality results in the most loss of information as it ignores whom the interactions are made with. Using distance from the event (i.e., last node in a walk) better associates agents to events but also ignores the specific interactions that result in such events. Finally, walk frequency considers all pairwise interactions that lead to specific events, and as we show later, is more representative of teamwork and can be used to predict the performance of teams more effectively.

Now, for each event, we will have different values for each metric for each agent or sets of agents (as for walk frequency). However, each event has a different impact on team performance (e.g., goals lead to a win, loss of possession likely to lead to a loss, person saved leads to successful rescue mission), and to determine the contribution of an agent or subset of agents to team performance, we need to learn the impact each metric has on the team’s performance. We assume that each event is independent⁴ and therefore use a weighted sum of the values for each of the possible events. This is shown in Equation 1.

$$v_m(a_i) = \sum_{k=1}^K \alpha_k v_m(a_i|e_k) \quad (1)$$

where, $v_m(a_i)$ is the value of a_i using the metric m , K is the number of possible events, α_k is the weight of the event e_k (which is learned from the data) and $v_m(a_i|e_k)$ is the value of a_i given the event e_k . We next expand on the above metrics in the following sub-sections.

Network Centrality Here we value an agent a_i based on its centrality in the network. This value is equal to the sum of the weights of the edges incident to node a_i (both incoming and outgoing edges). For example, in the graph shown in Figure 1, $v_{cent}(a_9|e_k) = w_8 + w_9 + w_{10}$. Equation 2

⁴In reality, some events may not be entirely independent and therefore, more complex summarisation functions would need to be used. But as we show in this paper, the assumption of independence does hold when it comes to predicting team performance.

shows the value calculation using the centrality metric for any agent a_i for the given event e_κ :

$$v_{cent}(a_i|e_\kappa) = \sum_{a_j \in \text{Adj}(a_i)} w(a_i, a_j) + w(a_j, a_i) \quad (2)$$

Distance From Event Given a set of events $\mathcal{E} = \{e_1, e_2, \dots, e_k\}$ and all the possible walks in graph G . The value for an agent a_i for any event $e_\kappa \in \mathcal{E}$ is defined as the average of the shortest path length of agent a_i from the event e_i for each walk where the agent a_i is present. The distance from the event of an agent a_i is the number of agents following a_i in a walk. For example, in Figure 1 in the walk $[a_4, a_9, a_5]$ the distance from event for a_9 is 2 and in the walk $[a_7, a_3, a_7, a_3, a_{11}, a_9]$ this distance for a_9 is 1. Hence, the average distance of agent a_9 from the event is $(1 + 2)/2 = 1.5$.

Walk Frequency The walk frequency of an ordered set of agents $A' = [a_x, a_y, \dots, a_z]$ based on the walk frequency is the number of times the agent A' appears in all the walks in G . For example, in Figure 1, suppose the walk $[a_4, a_9, a_5]$ appears three times, the walk $[a_7, a_3, a_7, a_3, a_{11}, a_9]$ appears four times, the walk $[a_3, a_{10}, a_8]$ appears twice and the walk $[a_1, a_6, a_2, a_3]$ appears once. In this case the value of $A' = [a_9]$ will be $v_{freq}(A'|e_k) = 3 + 4 = 7$. The same reasoning can be applied to subsets of agents.

We can compute such a metric for all sub-walks of a walk (i.e., subsets of a walk). Given a walk \mathcal{P} of length l , the number of sub-walks of length j constructed from the walk \mathcal{P} is calculated by picking the consecutive $j + 1$ vertices in the walk \mathcal{P} . The total number of sub-walks of a walk of length l is $\sum_{j=1}^{l-1} (l - j + 1)$ and extracting such sub-walks from each walk is relatively straightforward. As we show later in the paper (see Section 3.5), we focus on sub-walks involving pairs of agents as combine such pairs in a combinatorial optimisation algorithm to consider chains of interactions. We next describe how we will learn the weights of events α_k to compute Equation 1.

3.4 Learning Event Weights From Data

To learn the set of weights \mathcal{D} , which correspond to the impact of the possible walk events \mathcal{E} , we use a Logistic Regression algorithm (Hosmer Jr, Lemeshow, and Sturdivant 2013). This allows us to extract the coefficient weights of each of the input features i.e., the weight α_κ (which corresponds to an event e_i) which is used to calculate the final value $v_m(a_i)$ for each agent or sub-team of agents.

Hence, for an outcome y (e.g., a team wins a match, a political party wins an election), the probability that an agent a_i contributes to this outcome is dependent on the individual events (e_κ) to which an agent contributes, as captured by the metrics computed in the previous section. This can be summarised as per the linear combination in Equation 3.

$$p(y|a_i) = \alpha_0 + \alpha_1 v_m(a_i|e_1) + \dots + \alpha_k v_m(a_i|e_k) \quad (3)$$

The result of running the logistic regression algorithm is the set of weights $\alpha_\kappa \forall e_\kappa \in \mathcal{E}$. Given this, we can now compute efficient teams according to the learnt measures.

3.5 Forming Efficient Teams

We use two methods to form efficient teams using values calculated in the previous section. Firstly, we form teams based on the values of singleton agents. Secondly, we form teams based on the value of agent pairs p , so that teams are formed between agents who communicate and work well together.

Agent-Centred Approach To form the efficient team based on singleton agents, we use the values $v(a_i)$ for each agent a_i . Given constraints on the number of agents to be picked overall and the number of agents per role allowed in the team (see Section 5.2), this results in a combinatorial optimisation problem that is solved using standard mixed-integer programming (MIP) techniques. Similar methods are also used in (Pochet and Wolsey 2006; Fitzpatrick and Askin 2005; Matthews, Ramchurn, and Chalkiadakis 2012). Here we can use all the above metrics m (i.e., centrality, distance from event, and walk frequency).

Team-Centred Approach Here we consider how the team works effectively and hence only consider the walk frequency metric. Specifically, we reconstruct the value of teamwork based on two core concepts which we call the *strength of teamwork* and *interactional alignment* which we describe as follows.

- **Strength of Teamwork:** This is based on the contribution of the pairwise interactions, which in this case is shown by a high frequency of directed successful interactions between the agents. We calculate this in Section 3.3.
- **Interactional Alignment:** This is the measure of the strength of teamwork between overlapping pairs within the selected team. This values the strength of teamwork that the selected agents in a pair will bring when paired with other selected agents in the team. This helps us avoid selecting pairs of agents which have a strong value between themselves but are weak when combined with the rest of the team. We calculate this using Equation 5.

We combine these two measures to maximise the values of the selected pairs (p_i) while also maximising the value of the pairs that they overlap with in the selected team as a whole. Specifically, we propose an MIP defined by Equation 4. In more detail, the output of the MIP is a team of N agents formed by evaluating pairs of agents from the set $\mathcal{O} \subset 2^{\mathcal{A}}$ where for each $p \in \mathcal{O}, p \subset \mathcal{A}, |p| = 2$. We use two types of binary decision variables $x_i, z_j \in \{0, 1\}$ for pair i and agent j respectively. Variable x_i denotes whether a pair of agents is selected and z_j whether an agent is selected. The objective function maximises the sum of $V(p_i)$ (the value for p_i using the agent pair values we have calculated) and $V'(p_i)$ which represents the *interactional alignment* (the value of pair p_i calculated by Equation 5). This value is weighted by β which can be learnt from the data. The first constraint ensures that individual agents can be selected, even if they are in pairs that are not selected. The agent decision variables z_μ and z_λ represent the two agents in a given pair $p_i = \{a_\mu, a_\lambda\}$. The second constraint ensures that only N

agents are selected from all agents $\|\mathcal{A}\|$ available.

$$\begin{aligned} & \text{maximise} && \sum_{i=1}^{|\mathcal{O}|} \left(V(p_i) \cdot x_i + \beta V'(p_i) \cdot x_i \right) \\ & \text{subject to} && z_\mu \geq x_i, z_\lambda \geq x_i, \forall p_i = \{a_\mu, a_\lambda\} \\ & && \sum_{j=1}^{|\mathcal{A}|} z_j = N \end{aligned} \quad (4)$$

$V'(\cdot)$ is defined by Equation 5 as the sum of all pair values where there is an overlap with pair p_i . By maximising the *interactional alignment*, this allows us to increase the strong links between pairs while decreasing the weak links.

$$V'(p_i) = \sum_{k=1}^{|\mathcal{O}|} \left(V(p_k) \cdot x_k \right)_{\{p_i \cap p_k, k \neq i\}} \quad (5)$$

The generic solution presented in Equation 4 could also be expanded to consider the notion of roles within the team structure. An example of this would include the formation of a team in football where only one goalkeeper can be selected and the rest of the players are selected with different tactical roles (as we see later in Section 5.2). Similarly, within the emergency response domain we may need to form a team made up with specialists in different areas (e.g., paramedic and fire service). If roles are an important element to the team formation problem, we add extra constraints to our MIP formulation this is shown in Section 5.2.

4 Model Application

To validate the models defined in Section 3 we apply our techniques to the problem of team formation in football. In this section, we first give a background of the related work in football team formation and we then highlight how this relates to our model.

4.1 Football Related Work

Boon and Sierksma (2003) discuss the design of optimal teams and calculates the value-added from new team members. Following on from this, (Vilar et al. 2013) look to understand how players' and teams' strategies result in successful and unsuccessful relationships between teammates and opponents in the area of play. There have also been applications to form optimal teams for fantasy football using an MIP and performance predictions in (Matthews, Ramchurn, and Chalkiadakis 2012). Our models differ from the previous work as we model the team as a weighted-directed network of agents and value players based on their influence on the team and the teamwork between players. We then form teams using a novel algorithm with MIP techniques.

4.2 Football Application to The Model

In football, a manager/coach selects a team of 11 players from a squad of 23. The objective is to select a team with the highest chance of winning a game. Against this background, we define the squad of players as our set of agents \mathcal{A} , the interactions \mathcal{I} are the passes between the players in earlier

games, and the graph G represents the network of passes between all the players in the squad. The walk \mathcal{P} is a passage of play for the team which is made up from several passes. In football, a passage of play is ended by some event (e.g., tackle, shot, goal, miss and ball out of play). We characterise events into 4 possible outcomes, $\mathcal{E} = \{e_1, e_2, e_3, e_4\}$, where e_1 is a Goal, e_2 a shot on-target, e_3 a shot off-target and e_4 is a loss of possession. We then learn the weights α_i for each outcome. In this case we assume $\alpha_1 > \alpha_2 > \alpha_3 > \alpha_4$. Using the model discussed in the last section we calculate the value of each player $v(a_i)$ and form an optimal team based on the values considering the specific positional constraints of a football team. An example of a walk is shown in Figure 2 where the red arrows represent the passes between players and the blue arrow represents the outcome of the walk which in this case was a shot on target.

There are some positional constraints that are specific to football, making it more complex than the model we defined in Section 3.5. Each team in a game of football must have 1 goalkeeper and 10 outfield players which are formed from defenders, midfielders and strikers. In most positional formations in football, there are between 3-5 defenders, 3-5 midfielders and 1-3 strikers. An example formation is 4-4-2 which is 4 defenders, 4 midfielders and 2 strikers.

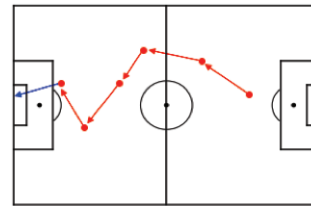


Figure 2: Example of a Walk in Football

5 Forming Efficient Football Teams

In this section, we describe the techniques that we use to solve our model and form efficient teams.

5.1 Calculating Player Values

To calculate the values of the players in the network, we first create the weighted graph that we need for our model. We do this using the walks (patterns of play) which happen in a game. We can then calculate their values for each of the possible walk events using each of the metrics defined in Section 3. The possible walk events we use are: a goal, a shot on target, a shot off target and a lost possession. Firstly, we do this for singleton players so that we have values based on their centrality, walk frequency and distance from the outcome. We then value the player pairs based on their frequency in the network. This gives us the values for both players and pairs from each match which we can then use to learn the impact weights of the outcomes.

Learning The Outcome Weights To calculate the weights of the walk events we use logistic regression as discussed in Section 3.4. Using the values for the players/pairs

for each walk event in each game we use the match outcome (team win, loss or draw) as the y value in our logistic regression formula. This means that we train the model to calculate the weights based on what impact it will have on the match outcome and, therefore, the overall team performance. The final value for the players/pairs will then be a weighted sum (defined in Equation 1) which uses these learned weights and will inform the team formation process.

5.2 Team Formation

We describe the two methods we take to form teams using both the singleton player values and the pair values.

Singleton Agents The first method uses the values of singleton players calculated using the centrality, walk frequency and distance from outcome (as discussed in Section 3). We use these values alongside constraints over players' positions to form the optimal team. The approach we use to solve this is an edited version of the MIP approach shown in Equation 6. Where we maximise $\sum_{i=1}^O (V(a_i) \cdot z_n)$ and do not consider the pair decision variable x_i . The other constraints remain the same.

Agent Pairs Using the values of the player pairs we form teams using the MIP formula presented in Equation 6 (this is a refinement of Equation 4). When forming teams we ensure that all the pairs of players are part of the same squad and can be selected together. We also consider the positions of the players so that we pick a team in a reasonable positional formation. This is represented by position range constraints.

$$\begin{aligned}
 & \text{maximise} && \sum_{i=1}^O (V(p_i) \cdot x_i + \beta V'(p_i) \cdot x_i) \\
 & \text{subject to} && \sum_{n=1}^{|A|} (z_n) = 11 \\
 & && z_\mu \geq x_i, z_\lambda \geq x_i, \forall p_i = \{a_\mu, a_\lambda\} \\
 & && \sum_{n=1}^{|A|} (gk_n \cdot z_n) = 1 \\
 & && 3 \leq \sum_{n=1}^{|A|} (def_n \cdot z_n) \leq 5 \\
 & && 3 \leq \sum_{n=1}^{|A|} (mid_n \cdot z_n) \leq 5 \\
 & && 1 \leq \sum_{n=1}^{|A|} (str_n \cdot z_n) \leq 3
 \end{aligned} \tag{6}$$

where a binary decision variable x_i represents the selected pairs ($p_i = \{a_\mu, a_\lambda\}$) and z_n represents whether a player is picked or not. There is then a number of binary sets for each position (gk , def , mid and str) containing if a player plays in the corresponding position or not and we aim to maximise the pair values V and V' in the selected team; β is set to 0.05.

6 Evaluation

To evaluate our models we use a dataset collected from the 2018 FIFA World Cup.⁵ The dataset breaks down the 64 games from the tournament into an event-by-event analysis where each event gives different metrics including: event type (e.g., pass, shot, tackle etc.), the pitch coordinates of the event and the event outcome. This type of dataset is industry-leading in football and used by top professional teams. Thus, we believe that this is a good, real-world, dataset with the

⁵All data provided by StatsBomb - www.statsbomb.com.

richness and challenge appropriate to rigorously assess the value of our model. To learn the model weights, we use a 10-fold cross-validation approach, splitting the dataset randomly into 70% training and 30% test.

6.1 Experiment 1: Performance Comparison to Human Formed Teams

We evaluate our model using all games from the 2018 FIFA World Cup and compare both the singleton approach and the pairs approach with the teams selected by the human-expert manager (focusing on both the starting 11 players and the 11 players who finish the game after substitutes). The results are presented in Figure 3 (where error bars represent a 95% confidence interval). The approaches used to form teams are:

- **Model 1:** Player values calculated using the centrality and team formed using an MIP.
- **Model 2:** Player values calculated using the walk frequency and team formed using an MIP.
- **Model 3:** Player values calculated using the distance from an event and team formed using an MIP.
- **Model 4:** Pair values are calculated and Equation 6 is used to form the team.

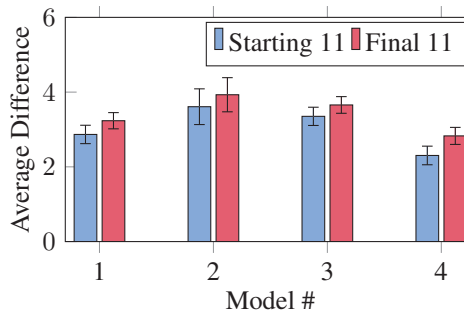


Figure 3: Average Difference Between Model and Real-World Human Manager Selections (where lower is better).

This shows that the pair values optimisation method gives the closest teams to the human experts on average with a difference of 2.3 per game for the starting team. This suggests that the human managers (either consciously or sub-consciously) consider the ability of players in the team to work together as the other methods only consider individual player values. At an average of 2.3, this could give managers suggestions of how changes could be made to the team that may give a better chance of winning the game.

6.2 Experiment 2: Outcome Prediction Comparison

We see that in Figure 4 there is a positive correlation (p -value = 0.001, r -value = 0.97) between the teamwork values and the number of goals scored by the teams and we see similar results for the correlation of the team value for other metrics. Hence, to evaluate the strength of our methods, we use the valuations as a predictor of the actual real-

world performance of the selected teams. We focus on match outcomes and other team performance metrics.

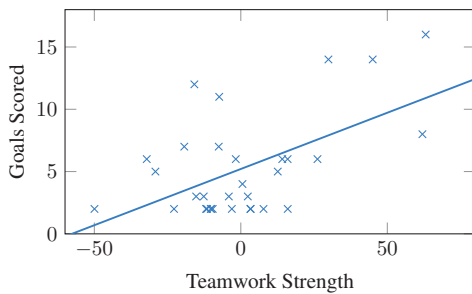


Figure 4: Correlation between team pair values and goals.

In Figure 5 and Table 1 we evaluate performance using a random forest approach to predict the team metrics (different to those in 4.2). These suggest that the teamwork metric is a more accurate predictor of performance than individual player values, meaning that the teams with higher valued pairs are more likely to win the game and have better performance indicators. This is especially true when we predict the number of passes that a team will make in a game as this metric shows the strongest predictor when using the teamwork values and is a 46% increase on any other approach.

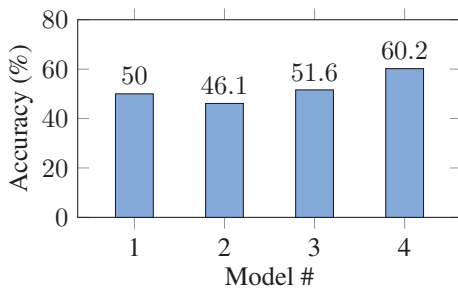


Figure 5: Accuracy of Valuation Methods For Outcome Predictions (Model 4 is the pairs value valuation).

Model #	1	2	3	4
Shots	4.33	4.52	4.33	3.74
Goals For	1.00	1.11	1.01	0.87
Goals Against	1.14	1.14	1.14	1.06
Passes	105.69	93.19	101.71	57.07

Table 1: Valuation Model Root Mean Squared Errors for Performance Prediction (where lower is better).

7 Discussion

We apply our model using the problems posed by football. However, there are many other domains where our models could be applied. The reason we have focused on football in this paper is that it presents a rich dataset that we can use to validate our model. There are also clear directed interactions

between players within a game of football (in the forms of passes) which allow us to form the networks to apply the proposed models.

We focus on forming teams based on singleton agent values and agent pair values. In further work, this could be extended to look at larger sub-team values and form teams based on those. However, we decided to focus on using the pair values to show the teamwork so that we can easily identify the pairs of agents that have a strong impact on team performance and the outcomes. This also allows us to calculate the *interactional alignment* of the pair so we can see how the selected agents will affect the rest of the team and therefore the overall teamwork of all agents.

We choose to test our team formation methods by comparing the outputs to that of a human-expert team manager who selects the real-world team. Our results show that our model can form teams which are similar to the selections of human-experts and, that we can suggest a small number of changes that could improve the team. This comparison also suggests that human-experts consider the teamwork between players when selecting their teams (this may be subconsciously).

Building on our models, in further work we would evaluate the performance of larger sub-team performance rather than singleton agents and pairs. We would further evaluate the predictions of match-outcomes, based on our team-valuations of a starting 11 team, against other match-outcome prediction approaches such as (Dixon and Coles 1997; Constantinou, Fenton, and Neil 2012). We would also extend the models to address how the team formation could be improved by factoring in an opposition team (in games such as football this can have a significant difference to how a team is formed). Our results also suggest that this model could be applicable across many domains and, given a high-quality dataset, we could further validate the model performance to see if similar results are found (e.g., in emergency response or data transfers).

8 Conclusion

In this paper, we describe a novel approach to team formation based on directed interactions between agents. Our model of teamwork considers event outcomes of the chains of interactions shown as walks within graphs. We defined and tested multiple network metrics to value the contribution of agents and sets of agents and show how the value of teamwork (including *interactional alignment*) can be learnt from data and then applied to predict the performance of teams. We tested and validated our models of valuing agents and forming teams by applying our models to problems posed by football and using data from the 2018 FIFA World Cup. We showed that our model is can produce similar team selections to an international level human-expert manager while also being suggesting changes to the team. We also showed how our valuation methods are an effective predictor of the key team performance metrics in football.

Acknowledgements

This research is sponsored by the EPSRC NPIF doctoral training grant number EP/S515590/1.

References

- Balkundi, P., and Kilduff, M. 2006. The ties that lead: A social network approach to leadership. *The Leadership Quarterly* 17(4):419–439.
- Bistaffa, F.; Farinelli, A.; Cerquides, J.; Rodríguez-Aguilar, J.; and Ramchurn, S. D. 2017a. Algorithms for graph-constrained coalition formation in the real world. *ACM Transactions on Intelligent Systems and Technology (TIST)* 8(4):60.
- Bistaffa, F.; Farinelli, A.; Chalkiadakis, G.; and Ramchurn, S. D. 2017b. A cooperative game-theoretic approach to the social ridesharing problem. *Artificial Intelligence* 246:86–117.
- Boon, B. H., and Sierksma, G. 2003. Team formation: Matching quality supply and quality demand. *European Journal of Operational Research* 148(2):277–292.
- Chalkiadakis, G., and Boutilier, C. 2012. Sequentially optimal repeated coalition formation under uncertainty. *Autonomous Agents and Multi-Agent Systems* 24(3):441–484.
- Constantinou, A. C.; Fenton, N.; and Neil, M. 2012. pi-football: a bayesian network model for forecasting association football match outcomes. *Knowledge-Based Systems* 36:322–339.
- Dixon, M. J., and Coles, S. G. 1997. Modelling association football scores and inefficiencies in the football betting market. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 46(2):265–280.
- Farinelli, A.; Iocchi, L.; and Nardi, D. 2004. Multirobot systems: a classification focused on coordination. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 34(5):2015–2028.
- Fitzpatrick, E. L., and Askin, R. G. 2005. Forming effective worker teams with multi-functional skill requirements. *Computers & Industrial Engineering* 48(3):593–608.
- Gaston, M. E., and DesJardins, M. 2005. Agent-organized networks for dynamic team formation. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, 230–237.
- Gaston, M. E., and DesJardins, M. 2008. The effect of network structure on dynamic team formation in multi-agent systems. *Computational Intelligence* 24(2):122–157.
- Hosmer Jr, D. W.; Lemeshow, S.; and Sturdivant, R. X. 2013. *Applied Logistic Regression*, volume 398.
- Matthews, T.; Ramchurn, S. D.; and Chalkiadakis, G. 2012. Competing with humans at fantasy football: Team formation in large partially-observable domains. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 1394–1400.
- Pochet, Y., and Wolsey, L. A. 2006. *Production planning by mixed integer programming*.
- Scerri, P.; Farinelli, A.; Okamoto, S.; and Tambe, M. 2005. Allocating tasks in extreme teams. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, 727–734.
- Vilar, L.; Araújo, D.; Davids, K.; and Bar-Yam, Y. 2013. Science of winning soccer: Emergent pattern-forming dynamics in association football. *Journal of Systems Science and Complexity* 26(1):73–84.
- Zhong, X.; Huang, Q.; Davison, R. M.; Yang, X.; and Chen, H. 2012. Empowering teams through social network ties. *International Journal of Information Management* 32(3):209–220.