Fantasy Soccer as a Contextual Adversarial Bandit Problem: Algorithmic Approaches to Strategic Team Selection and 'On-Field Decision-Making'

Faris Qadan
School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, USA
fqadan3@gatech.edu

December 14, 2023

Abstract

Fantasy soccer (FS) is one of the most widely applicable instances of an accessible online decision-making problem. As such, this project poses FS as an online bandit problem. We explore this outline from both an adversarial and a contextual lens, and empirically test to see which formulation is more suitable in playing a specific part of the game. Finally, we show that by its current definition, FS is very suitable for analysis as a contextual bandit, and we suggest the modifications that may be required to view it more adversarially.

1 Background

1.1 Brief Background on Fantasy Soccer

Fantasy soccer (FS) involves users selecting 15 soccer *players* to build a team based on a fixed budget. Note that 'user' and 'player' are NOT used interchangeably in this project, given their different roles in this problem. One hour before the soccer games commence for the gameweek (in real-life), users' teams cannot be further modified. They then observe the total number of points they accumulated for the week after all the games are played and completed.

Players earn points based on their contributions to real life soccer games every gameweek. For example, a goal scored earns a player 5 points, while a goal conceded loses them 5 points (or more precisely, earns them -5 points). Users must also select one captain every week. This pick is vital, as a user's selected captain earns twice the number of points than they normally would (i.e. their regularly earned points are twice as heavy).

In short, the user's goal is to achieve the maximum number of points over the course of a 38-week long season, given the constraints mentioned above (and some others, omitted here for brevity). To provide a point of reference for a team's accumulated points, the top FS finishers over the last 11 years have averaged just over 2600 points (around 68 points per week), while my personal teams have averaged close to 1400 points per year (37 points per week).

1.2 Problem Formulation

We consider a setup where each gameweek, g, models a single time-step or decision round, after which we observe a reward, p. We are only able to refine our selection strategy based on our reward and given contexts. As such, it is clear that FS is an online decision-making problem.

In this project, we will focus on a very specific decision in the FS problem: captain selection. Captain-acquired points account for 20-30 % of a user's total point tally. Hence, a focus on this decision can be thought of as a condensed, more tractable model of the fantasy soccer problem.

1.3 Adversarial and Contextual Bandits

In brief, adversarial and contextual bandits are primarily seperated by their objectives: adversarial bandits aim to 'beat' an adversary, while contextual bandits more simply look to maximize a wealth of context in their decision-making.

In the context of fantasy soccer, context is valuable and plentifully available. This includes player-based context (such as goals scored, 'expected goals', and more) and team-based context (such as fixture difficulty rankings and home or away advantages). It is thus clear that the problem can (and likely should) be modelled as a contextual bandit.

Less intuitively, however, is the potentially adversarial nature of the FS games. One nicely translatable example of this in my experience of playing FS is as follows: Pep Guardiola, team manager of Manchester City, is notorious for his heavy team rotation. Manchester City has won 5 of the 6 last titles, with many of their players performing exceptionally along the way. However, given Guardiola's heavy rotation of these players, it has often not been in one's best interest to choose a Manchester City player as their captain. Pep Guardiola could be considered 'adverserial' here. Both perspectives are thoroughly explored throughout the rest of this report.

1.4 Related Works

To the best of my knowledge, fantasy soccer has never been explicitly posed as a bandit problem. However, there have been some attempts in 'automating' the FS experience. One such effort by Anand [2022] involves using regression analysis to eliminate favoritism bias in team selection (a factor that I argue is adversarial). Most other efforts have been along these lines.

Most attempts found use a form of time-series analysis, combined with regression, to create a prediction model of sorts.

Finally, some instances were found to use Q-learning in selecting a starting lineup for a given fantasy team. Although this aligns more closely to this task at hand, the literature that proposed these methods was not as reputable as that for the methods above and mostly came in the form of recreational explorations from hobbyists. As such, the Q-learning-based literature was not used in reference to this.

1.5 Novelty and Approach

This project explores adversarial and contextual bandit algorithms as they pertain to solving the problem of captain selection. More specifically, we aim to use our results to evaluate whether FS can be modelled as an adversarial bandit, to compare its modelling as an adversarial and a contextual bandit, and to arrive at conclusions and suggestions for future bandit-based modelling of the FS problem.

The basis of this approach will include two slightly-tweaked implementations of Exp3 [Auer et al., 2003] and LinUCB [Li et al., 2010] that will be detailed and evaluated below.

2 Methodology

Given the stated research goals for this project, we begin by modelling FS adversarially and contextually. To further outline this problem, at some round 't', we will run our algorithm in order to select the best player (arm), 'A_t' from a list of arms.

One constraint, however, adds novelty to this problem as compared to the 'vanilla' bandit problem: on any given round, it is possible that the list of k possible captains (k-arms) differs from those before it. This is to reflect the true nature of fantasy soccer: it is a dynamic game where often, players experience sudden surges and drops in form unbeknownst to users. As such, we present FS-Exp3 and FS-LinUCB – two slightly-tweaked algorithms to accomodate for this constraint.

2.1 Dataset

The dataset we use was generously compiled and mostly cleaned by Anand [2022]. It is publicly available on GitHub, and contains a list of all the players in the English Premier League (upon which our version of FS is based on), alongside relevant statistics and metrics that will be used as context in later portions of this project.

For the pursposes of this work, we used Python to further clean and process this data. More specifically, we made player formats more readable, and wrote functions to easily access relevant information for the below sections of the report.

2.2 Arm Filtering

At any given round, there are hundreds of players available for selection as an arm. In order to make this problem more computationally feasible, we obtain a list of 6 arms every round based on the following criteria:

- Player cannot be a goalkeeper (historically the lowest scoring position)
- Players must be in form (at least 6.5 points per game over the last three games)
- Players must play consistently (at least the equivalent of 30 out of the last 38 games)

In the years for which data is available, this list has never been sized below six players. In the rare instances that this list of arms exceeds a size of six, the six arms are chosen from a uniform distribution to be the candidate arms for the round.

2.3 FS-Exp3

As defined by Auer et al. [2003], Exp3 is a policy-based algorithm that involves selecting an arm by sampling from the list of arms, based on an updating probabalistic set of weights (P_t is the set of weights for all arms at round t).

In order to account for the notion of a dynamic list of arms, we introduce the notion of a weight-map. More specifically, we keep a map with all the arm weights accumulated up to some round t. At t + 1, we devise our list of six arms, and obtain their relevant weights. We then normalize these weights, such that based on the criteria from which our list of arms was chosen, we are pulling the 'best' arm locally. If a new, unseen arm is added to a list, they are given a weight from a normal distribution, and the local arm list is normalized before proceeding with the same process.

Finally, upon selection of an arm based on the Auer et al. [2003] version of Exp3, the local weights are put back into the broader, global weight map and renormalized.

To the best of my knowledge, no scheme of this sort (at least for this application) has been used before. However, this scheme allowed for two primary points to be accomplished: (a) selection of the best captain in a local domain, based on the filtering criteria above, and (b) allowing for round-by-round changes to affect change on the broader season-long landscape. In other words, we first select the best arm locally, before noting this update globally.

The below algorithm is based on the version of Exp3 outlined by Lattimore and Szepesvari [2017] in their book, *Bandit Algorithms*.

Algorithm 1 FS-Exp3

Require: η, gameweek

$\hat{S}_0(i) := 0$ for all i and weight map := {}
for $t = 1, \ldots, n$ **do**
    Generate arms list
    for each arm a in arms **do**
        if a not in weight map **then**
            $P_{t,a} \sim N(0,1)$
            Add a to weight map, renormalize
        end if
    end for
    Calculate the sampling distribution P_t:

$$P_{t,i} = \frac{\exp(\eta \hat{S}_{t-1,i})}{\sum_{j=1}^{k} \exp(\eta \hat{S}_{t-1,j})}$$

    Sample $A_t \sim P_t$ and observe reward X_t
    Update local weights $\hat{S}_t$:

$$\hat{S}_{t,i} = \hat{S}_{t-1,i} + 1 - \frac{I\{A_t = i\}(1 - X_t)}{P_{t,i}}$$

    Update global weight map with local weights and renormalize
end for

2.4 FS-LinUCB

LinUCB is a contextual bandit algorithm that takes an optimistic approach on selecting an arm per round. Specifically, we build a confidence set containing the optimal arm and pull the arm with the highest confidence bound each round, updating our 'level of confidence' as we observe our reward. In a conventional setting, LinUCB offers an efficient, quick way of converging to a typically optimal arm.

In this context, we face the same problem here as in the adversarial setting. To accommodate for this, we develop the notion of a 'dynamic Gram map', whereby the feature vectors for each selected arm are chosen for the Gram matrix at a specific round. That is, in a similar way to above, we first make a local decision, before updating our confidence on a global scale.

The feature vectors here will be detailed with context from the aforementioned database. That is, in one instance, context for a certain player A can be given by the number of goals, assists, and total points they scored in round $t - 1$.

This is outlined in the algorithm below. This implementation of LinUCB is based on that shown in Lecture 18 of this course's notes.

Algorithm 2 FS-LinUCB

Require: $\delta, A,$ gameweek

$\hat{\theta}_0 \in R^d := 0$
for $t = 1, 2, \ldots$ **do**
  Generate arms list
  for each arm a **do**
    if a is not in list of all arms **then**
      Update G_t
      Get a's context
    end if
  end for
  Estimate $\hat{\theta}_t$
  Select action $A_t = \arg\max_{a \in A} \max_{\theta \in C_t} < a, \theta >$
  Observe reward for the action taken
  Update G_t and θ_t and renormalize
end for

To summarize, this is very similar to the original rendition of the LinUCB algorithm, but with accomodations for the dynamic list of arms.

2.5 Experimental Setup

To test our research objectives, we set up two individual experiments. The first utilizes the Exp3 algorithm above. Specifically, we test the FS-Exp3 algorithm over a random season for which data is available. To better understand the intricacies of our modelling of FS as an adversarial bandit, we vary the learning rate from close to 0, up to close to 1.

This will allow us to come to multiple different conclusions, depending on the result. Most importantly, though, it will allow us to understand the extent to which this problem can be modelled as adversarial. The better our algorithm performs with a lower learning rate (and more randomness), the harder it is for a hypothetical adversary to 'play against us', and vice-versa.

With regards to the contextual bandit setting, we look to explore the extent to which context is useful (or necessary) for converging to the best arm quickly. As such, we perform a similar simulation with FS-LinUCB to the above, with varying amounts of context (and thus, varying dimensions of any given G instance).

This will allow us to determine whether extra context is always good. Specifically, we look for whether any diminishing returns are experienced by increasing the amount of context (and thus, computational time and power required) in solving the problem.

In both cases, we use the real best captain choice (in hindsight) as our benchmark. To compare our algorithms to this, we model the disparity in cumulative points they experience as compared to the benchmark. That is, we sum the total difference in points between the best possible choice and ours over 'T' rounds.

3 Results and Discussion

As stated above, we performed both experiments on a randomly selected English Premier League season, whose data is available in the aforementioned dataset. Below, we show the cumulative point discrepancy between FS-Exp3 with a learning rate of 0.1 and one of 0.65.

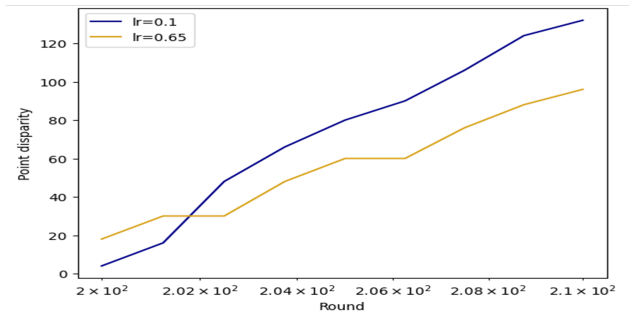

Figure 1: Cumulative point disparity between FS-Exp3 w/ LR=0.1 and 0.6

There are two interesting takeaways from the results we see in Figure 1. Firstly, we see that in both cases, we experience a somewhat linear growth in point disparity. At no point do we converge towards a single predictable solution. This is not unexpected: given the nature of Exp3 (and the fact that it is designed to play against an adversary), this behavior is desirable.

Furthermore, we see in Figure 1 that FS-Exp3 with a very low learning rate performs terribly (as compared to the more moderately chosen one). This means that our algorithm heavily benefits from added history, and suffers heavily when it 'gambles' or picks more erratically.

In tandem, these two points lead to a similar conclusion: as we have currently chosen to model it, fantasy soccer is not suited for analysis and resolution in an adversarial setting. This can be for many reasons. For one, the adversary (or adversaries) in this setting are not clearly defined. While some factors could be argued to be adversarial in nature, FS models appear to follow a certain predictability over time. That is, since the adversary is not clearly defined in this simplified model and typically, the number of rounds these algorithms are being performed on is relatively short in nature, we benefit heavily from exploiting past data and refraining from very heavily selecting random arms.

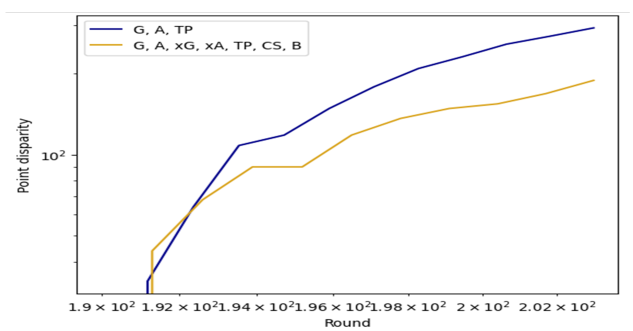

Figure 2: Cumulative point disparity between FS-LinUCB w/ 3 vs. 6 context points

Above, we see the results for the performance of FS-LinUCB with varying levels of context over the same span of games shown above. It should first be noted that, contrary to our application of Exp3's results shown above, we notice a level of convergence in both cases. This, coupled with the deductions made above, reaffirm our definition of the FS problem as a contextual bandit.

More curiously, we note that a larger set of context used yields convergence to a better arm, at an earlier stage. Looking more in depth, we see that despite the convergence of both instances, the convergence actually occurs to two different players (arms): the 3-point algorithm converges to Patrick Bamford (an attacker), while the 6-point algorithm converges to Bruno Fernandes (a midfielder).

FS-wise, midfielders are rewarded more heavily for more factors than attackers, such as clean sheets and goals. We further note that Bruno Fernandes was actually the optimal arm (highest scoring player) throughout the season around which this simulation ran. As such, the additional context (including these two pieces of information) give our algorithm a more comprehensive picture of the arms' landscape, allowing it to converge to a better arm overall.

While the benefit of the added context is not unexpected, the stark performance difference between the two was unexpected. Although it was clear than LinUCB (or some variation of it) should converge to an arm, its convergence to a somewhat far suboptimal arm was a surprise.

4 Conclusions

Overall, our results sufficed in helping answer the research objectives posed initially. Firstly, we conclude that as it is currently set up, our model of FS is not suited for analysis as an adversarial bandit. In fact, when tackling it with an algorithm catered to adversarial bandits, the better solutions came as we strayed further from treating the problem adversarially, and closer to treating it contextually.

To model this adversarially in a more effective light, our definition of what an adversary is in this context would have to be refined. In this model of FS, it is clear that excessive randomization actually hurts the performance of a given algorithm due to this.

With regards to its posing as a contextual bandit, it was shown that the richer the context used, the better our contextual bandit algorithm performed in solving the FS problem. In fact, a rich set of context enable our algorithm not only to converge to the best solution, but to do so slightly earlier than the case with less context.

As such, we can say that in this current, simplified setup, the best option is to pose this problem as a contextual bandit. The biggest bottleneck in the favorable contextual solution is computation: the richer the context, the more computation is required to make a decision per round. Although hypothetically, these rounds occur sporadically, this solution is likely not scalable.

Hence, a compromise to solving this problem more tractably would be to explore which pieces of context are the best indicators for player performance over time. That is, what is the least amount of information we need to make a good choice most of the time?

In conclusion, by implementing novel adaptions of Exp3 and LinUCB, we saw that the FS problem defined is most suited for modelling as a contextual bandit. Our FS-Exp3 algorithm performed suboptimally in multiple cases, but left us with very valuable conclusions. On the other hand, our FS-LinUCB algorithm with 6 points of context yielded impressive results when compared to the benchmark in retrospect, leaving room for development in this regard during future work.

5 Future Work

As mentioned in depth above, our contextual bandit model for fantasy soccer was highly successful. Hence, the first course of action in future work would be to more extensively test the algorithm out. No multi-season tests were done during these months due to the massive time and computational resources required to do so. This was the primary cause in the limited amount of results obtained – the more time and computational power is available, the closer we can get towards robust conclusions with regards to our initially-stated research objectives.

Following on from this, refining the algorithms would also be vital. While they are sequential in nature, finding better ways to filter arms, select them, and handle the dynamic set of arms will be key in scaling this project to a successful, winning fantasy player. With this, regret analysis would also be performed on the refined algorithms, ensuring more formal definitions and expectations for our results as compared to the best possible performances in hindsight.

Furthermore, as mentioned previously, more in-depth explorations into the most important pieces of context for point-prediction in FS will also be very important. As efficient as the algorithms themselves get, work with matrices is notoriously resource-intensive. As such, the more information we can extract from less-rich matrices, the better.

Finally, an exploration into a better way to model adversaries in this problem's context would be welcome. While it is clear that *context* is vital in effectively solving the FS problem, we often witness cases of previously extremely suboptimal captain picks unexpectedly over-performing or reaching surges of form. Our models' tendencies against randomness prevents us from exploiting situations like this. This could be one way of modelling adversaries, and would be a highly interesting avenue to take the project towards in the future.

With this, we would look into exploring a joint algorithm, ideally combining the best of both modelling aspects of the problem. In theory, this would allow for the most comprehensive model of FS given the plethora of historical data available.

References

Vaastav Anand. FPL Historical Dataset. Retrieved August 2022 from `https://github.com/vaastav/Fantasy-Premier-League/`, 2022.

Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The nonstochastic multiarmed bandit problem. *SIAM J. Comput.*, 32(1):48–77, jan 2003. ISSN 0097-5397. doi: 10.1137/S0097539701398375. URL `https://doi.org/10.1137/S0097539701398375`.

Tor Lattimore and Csaba Szepesvari. Bandit algorithms. 2017. URL `https://tor-lattimore.com/downloads/book/book.pdf`.

Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, page 661–670, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781605587998. doi: 10.1145/1772690.1772758. URL `https://doi.org/10.1145/1772690.1772758`.