Using PageRank to Predict the Netball

Olivia Watson   Harry Spearing

September 5, 2019

Introduction

The Aim:

- To develop an accurate and robust ranking system that is applicable for all sports, before then tailoring it to a specific sport to improve its accuracy.
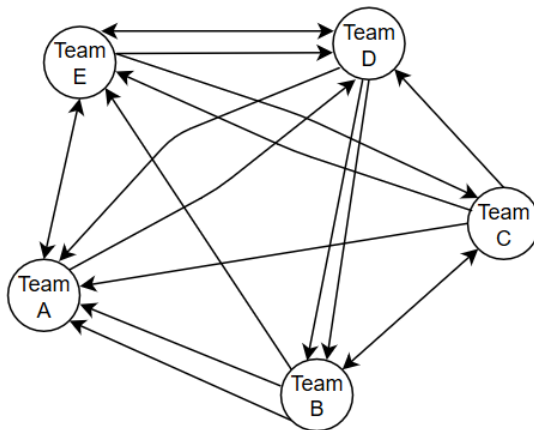
Constructing the Model

Figure: Network Example

Constructing the Model

- In the base model adjacency matrix, V, each team is awarded a value of α (where $\alpha \in (0.5, 1)$) for each match that they win, a value of 0.5 for each match they draw and a value of $1 - \alpha$ for each match they lose.

Constructing the Model

- In the base model adjacency matrix, V, each team is awarded a value of α (where $\alpha \in (0.5, 1)$) for each match that they win, a value of 0.5 for each match they draw and a value of $1 - \alpha$ for each match they lose.

$$V = \begin{bmatrix} 0 & 0 & 0 & 1 & \frac{1}{2} \\ 2 & 0 & \frac{1}{2} & 0 & 1 \\ 1 & \frac{1}{2} & 0 & 1 & 1 \\ 1 & 2 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & \frac{3}{2} & 0 \end{bmatrix} \qquad P = \begin{bmatrix} 0 & 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{4}{7} & 0 & \frac{1}{7} & 0 & \frac{2}{7} \\ \frac{2}{7} & \frac{1}{7} & 0 & \frac{2}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{4}{7} & 0 & 0 & \frac{1}{7} \\ \frac{1}{6} & 0 & \frac{1}{3} & \frac{1}{2} & 0 \end{bmatrix}$$

Solving the Model

- Mathematically, given the transition probability matrix P, the rating of team j can be written as:

$$\pi_j = \sum_{i=1}^{n} p_{ij} \pi_i$$

Idea Refinement

To develop the PageRank sport specific features relating to netball were looked at, such as

- Home advantage
- Incentive to win
- Player injuries
- Margin of victory
- Winning streak

Idea Refinement

To incorporate margin of victory, for example, an indicator function was used in the base model adjacency matrix.

$$\lfloor(|G1 - G2| - \delta)\beta\rfloor + \zeta + 1$$

where goal difference is $|G1 - G2|$, δ is the lower bound for the range of goals, β is the weight given to that range of scores, ζ is the number of points already awarded prior to that goal range.

For $10 \leq |G1 - G2| \leq 26,$ $\quad \beta = \frac{1}{4}$ $\delta = 10$ $\zeta = 1$

For $26 < |G1 - G2| \leq 32,$ $\quad \beta = \frac{1}{3}$ $\delta = 26$ $\zeta = 5$

For $32 < |G1 - G2|,$ $\quad \beta = \frac{1}{2}$ $\delta = 32$ $\zeta = 7$

Results

	Home Team Wins	Bookmakers	PageRank
National League	0.417	0.454	0.424
Premier League	0.476	0.584	0.524

Table: Accuracy of Football Predictions

	Vitality, 2018	Vitality, 2019	Australia, 2019
Without	0.622	0.667	0.577
With	0.656	0.656	0.596

Table: Accuracy of PageRank for Netball

Compared to the Bookmakers

	Round 13	Round 14
Bookmakers Prediction	Swifts	Vixens
	Giants	Giants
	Fever	Swifts
	Draw	Lightning
PageRanks Prediction	Swifts	Magpies
	Giants	Giants
	Fever	Swifts
	Lightning	Lightning
Winner	Neither	PageRank
	Both	Both
	Neither	Both
	PageRank	Both

Table: Season Comparison to the Bookmakers

Conclusions and Future Enhancements

Conclusion

- It appears that the bookmakers have many more variables being taken into account for football data then they do for netball data.

Future Enhancements

- Currently, the PageRank algorithm is only designed to work for one-on-one matches but I would like to adapt it such that it can be applicable to sports that are not one-on-one.

Any Questions?