Using PageRank to Predict the Netball

Olivia Watson          Harry Spearing

September 5, 2019
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 $\alpha$ (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 $\alpha$ (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.

$$V = \begin{bmatrix}
0 & 0 & 0 & 1 & \frac{1}{2} \\
2 & 0 & \frac{1}{2} & 0 & 1 \\
1 & \frac{1}{2} & 0 & 1 & 1 \\
1 & 2 & 0 & 0 & \frac{1}{2} \\
\frac{1}{2} & 0 & 1 & \frac{3}{2} & 0
\end{bmatrix} \quad P = \begin{bmatrix}
0 & 0 & 0 & \frac{2}{3} & \frac{1}{3} \\
\frac{4}{7} & 0 & \frac{1}{7} & 0 & \frac{2}{7} \\
\frac{2}{7} & \frac{1}{7} & 0 & \frac{2}{7} & \frac{2}{7} \\
\frac{2}{7} & \frac{4}{7} & 0 & 0 & \frac{1}{7} \\
\frac{1}{6} & 0 & \frac{1}{3} & \frac{1}{2} & 0
\end{bmatrix}$$
Mathematically, given the transition probability matrix $P$, the rating of team $j$ can be written as:

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

$$\left\lfloor (|G_1 - G_2| - \delta) \beta \right\rfloor + \zeta + 1$$

where goal difference is $|G_1 - G_2|$, $\delta$ is the lower bound for the range of goals, $\beta$ is the weight given to that range of scores, $\zeta$ is the number of points already awarded prior to that goal range.

For $10 \leq |G_1 - G_2| \leq 26$, $\beta = \frac{1}{4}$, $\delta = 10$, $\zeta = 1$

For $26 < |G_1 - G_2| \leq 32$, $\beta = \frac{1}{3}$, $\delta = 26$, $\zeta = 5$

For $32 < |G_1 - G_2|$, $\beta = \frac{1}{2}$, $\delta = 32$, $\zeta = 7$
### Results

<table>
<thead>
<tr>
<th></th>
<th>Home Team Wins</th>
<th>Bookmakers</th>
<th>PageRank</th>
</tr>
</thead>
<tbody>
<tr>
<td>National League</td>
<td>0.417</td>
<td>0.454</td>
<td>0.424</td>
</tr>
<tr>
<td>Premier League</td>
<td>0.476</td>
<td>0.584</td>
<td>0.524</td>
</tr>
</tbody>
</table>

**Table: Accuracy of Football Predictions**

<table>
<thead>
<tr>
<th></th>
<th>Vitality, 2018</th>
<th>Vitality, 2019</th>
<th>Australia, 2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without</td>
<td>0.622</td>
<td>0.667</td>
<td>0.577</td>
</tr>
<tr>
<td>With</td>
<td>0.656</td>
<td>0.656</td>
<td>0.596</td>
</tr>
</tbody>
</table>

**Table: Accuracy of PageRank for Netball**
## Compared to the Bookmakers

<table>
<thead>
<tr>
<th></th>
<th>Round 13</th>
<th>Round 14</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bookmakers Prediction</strong></td>
<td>Swifts</td>
<td>Vixens</td>
</tr>
<tr>
<td></td>
<td>Giants</td>
<td>Giants</td>
</tr>
<tr>
<td></td>
<td>Fever</td>
<td>Swifts</td>
</tr>
<tr>
<td><strong>PageRanks Prediction</strong></td>
<td>Swifts</td>
<td>Magpies</td>
</tr>
<tr>
<td></td>
<td>Giants</td>
<td>Giants</td>
</tr>
<tr>
<td></td>
<td>Fever</td>
<td>Swifts</td>
</tr>
<tr>
<td><strong>Winner</strong></td>
<td>Neither</td>
<td>PageRank</td>
</tr>
<tr>
<td></td>
<td>Both</td>
<td>Both</td>
</tr>
<tr>
<td></td>
<td>Neither</td>
<td>Both</td>
</tr>
<tr>
<td></td>
<td>PageRank</td>
<td>Both</td>
</tr>
</tbody>
</table>

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 than 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?