Coach Ranking Model

Akilesh Potti, Dominick Twitty, Wenhai Yang

February 8, 2014

Abstract

Todo

Disclaimer: For legal and moral reasons, all geographical locations and company names used in this paper are entirely fictional unless otherwise stated. However, for the sake of accuracy, real world data was used to tune our model.

1 The Problem

1. Todo

2 Model assumptions

1. Todo

3 Models

3.1 Simple Heuristic Model

3.1.1 Sorting by Win/Loss Ratio

3.1.2 Sorting by Net Wins

3.1.3 Differentiate between Games

3.2 Machine Learning Model

3.3 Graphical Model

It comes very natural for us to consider the problem as a graph. The nodes of the graph are the coaches. An edge between two nodes denotes the relation between two coaches, such as the game result of them playing against each other.

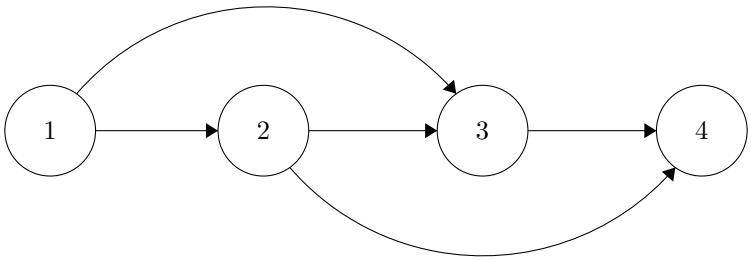
This idea of representing the problem as a graph is a very good step forward because it allow us to go beyond the simple statistics like win/loss ratio of a single coach, and gain more insight from the data.

Another advantage of this model is that it can aggregate data from different time periods. For example if coach A beats coach B and then retired in 1990, and coach B beat coach C later in 1995, then in the graph we will have a way to compare coach A and coach C because there is a path from A to C through B, even though they never played against each other.

3.3.1 Topolical Ordering

A very first idea that comes to mind is Topological Ordering. In this sub-model, there exist an edge from i to j if and only if among all the games i played against j, i beat j more often than j beat i. Then if there exists a Topological Ordering of the graph, we have a ranking of the coaches based on their rival history.

Consider the following example. Coach 1 has been beaten by both 2 and 3, coach 2 has been beaten by coach 3 and 4, coach 3 has been beaten by coach 4, and coach 4 is never beaten by anyone. We have a topological ordering in the graph: $\{1, 2, 3, 4\}$, and it is natural to conclude that coach 4 is the best coach among them.



However, one natural limitation of this model is the following theorem:

Theorem 3.1 *A graph G has a topological ordering if and only if it is a DAG (directed acyclic graph).*

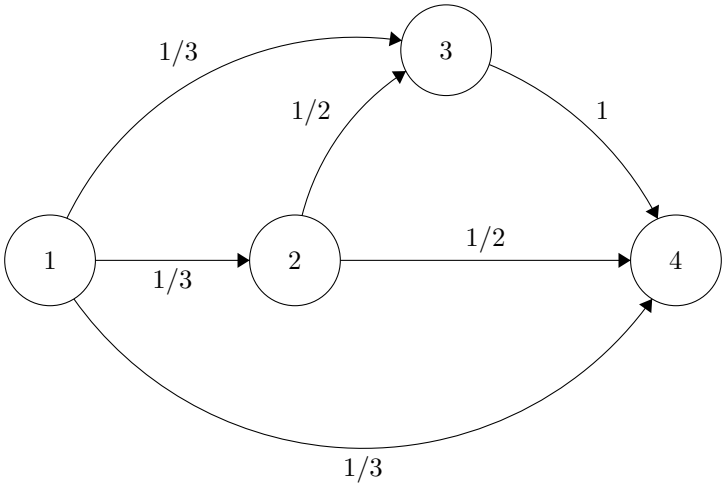
Proof See Section 3.6 of Algorithms Design by Kleinberg and Tardos.

To solve this issue, we think of an algorithm to reduce the originial graph into a DAG:

Algorithm 1: Graph Reduction Algorithm

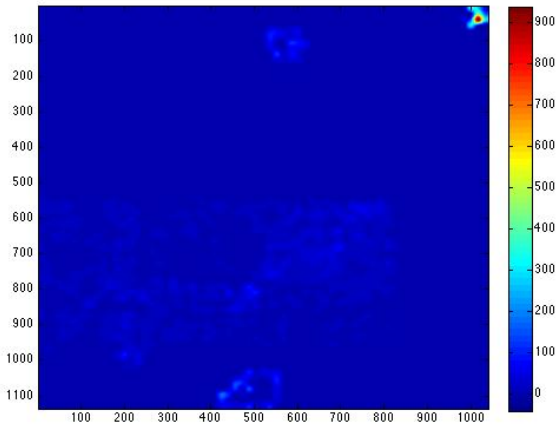
while *G is not a DAG* **do**
  e = edge with the least importance in G;
  Remove e from G
end

3.3.2 Markov Chain

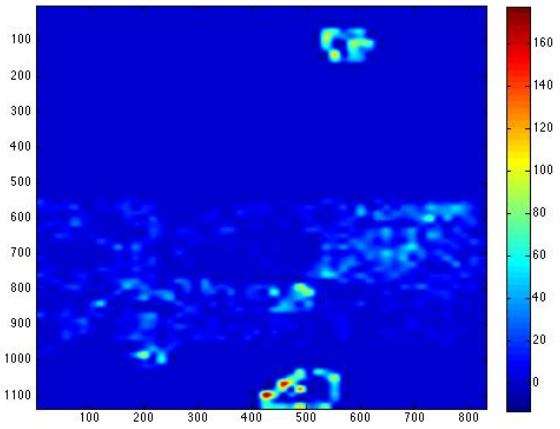


4 Results, Validation, and Robustness

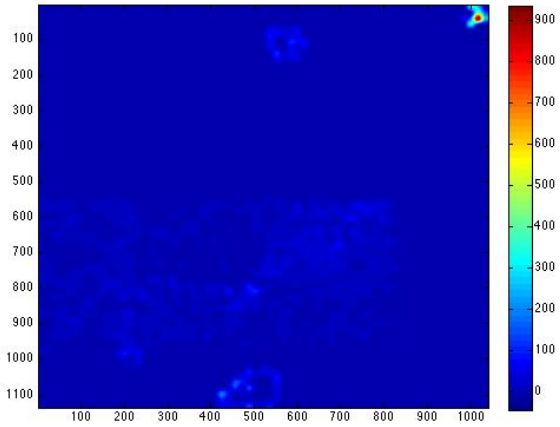
(a) Distribution of Request Source with Airport



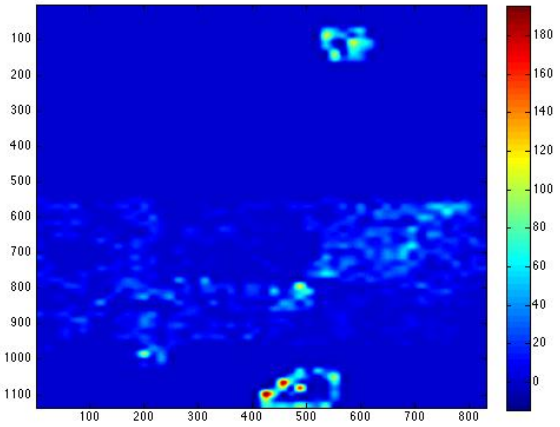
(b) Distribution of Request Source without Airport



(c) Distribution of Request Destination with Airport



(d) Distribution of Request Destination without Airport



$$p_{unreg} = k \ln(t_{reg})$$

$$\frac{p_{unreg'}}{p_{unreg}} = \frac{\ln_{t_{reg}}}{\ln t_{reg'}}$$

Number of Cabs	14	15	16	17	18	19	20	21
Uni Ind/Conglo	1.5720	1.3773	1.2880	1.2071	1.1932	1.2113	1.1201	1.0672
Prop Ind/Conglo	1.6876	1.4040	1.2459	1.1248	1.2558	1.1498	1.1033	1.1022
Disprop Ind/Conglo	2.1519	2.0706	2.0489	1.9399	1.8725	1.8174	1.6990	1.3905

5 Strengths and Weaknesses

5.1 Strengths

- Todo

5.2 Weaknesses

- Todo

6 Conclusions

Todo

7 Future work

- Todo

8 Bibliography

References

[1] OpenStreetMaps API, http://www.openstreetmap.org/#map=14/42.4427/-76.4984

[2] Github Gist by user aflaxman, https://gist.github.com/aflaxman/287370/

[3] A stochastic optimization model for real-time ambulance redeployment, http://www.sciencedirect.com/science/article/pii/S0305054813000385

[4] Ithaca Demographics, http://www.ci.ithaca.ny.us/maps/index.cfm

[5] Cornell Demographics, http://www.cornell.edu/about/facts/stats.cfm

[6] Ithaca College, http://www.ithaca.edu/admission/facts/

9 Code