Voter model on NBA ballots for Most Valuable Player

Faglioni Isacco
Department of Physics,
Modena, Italy email:
isacco.faglioni@studio.unibo.it

w

Keywords: Complex Networks, Complex System, Graph Theory

1 Introduction

Complex Network Analysis (CNA) has evolved as a powerful tool to comprehend, visualize, and quantify networks, offering insights into their inherent characteristics and the underlying principles dictating their formation and growth. Through the study of nodes (the entities) and edges (their interactions), CNA uncovers hidden patterns, identifies pivotal entities, and predicts network behavior, among other capabilities.

This article focuses on bipartite networks and the so called voter model. Bipartite networks constitute a fundamental framework for modeling complex relationships in a wide array of natural and artificial systems. Unlike traditional networks, bipartite graphs are characterized by two distinct sets of nodes, where connections exclusively occur between nodes from different sets. This structure finds applications in diverse domains including social interactions, ecological networks, and recommendation systems. Within this context, the Voter Model stands as a cornerstone in understanding the dynamics of opinion formation and adoption in bipartite networks. Originating from statistical physics, the Voter Model serves as a mathematical representation of how individuals alter their preferences or opinions based on the influence of their network neighbors. In bipartite networks, this model provides a powerful tool for studying the propagation of preferences among distinct classes of nodes, offering insights into phenomena such as information diffusion and consensus emergence.

1.1 Voter model on NBA ballots. In this article, we explore the dynamic of consensus in the NBA voting for Most Valuable Player. Through empirical studies and computational simulations, we aim establish if the model is fit for making predictions on small complex networks.

2 Bipartite Graphs: A Foundation for Modeling Complex Systems

Bipartite graphs, a distinctive class of networks, provide a powerful framework for representing and analyzing complex relationships. Unlike conventional graphs, bipartite structures explicitly distinguish between two disjoint sets of nodes, such that edges exclusively connect nodes from different sets. This peculiarity makes them a perfectly fitted tool to study systems that can be modelized with interactions between disjoint sets such as an election.

2.1 Structure and Formalism. Formally, a bipartite graph $G = (U, V, E)$ consists of two distinct sets of nodes, U and V, with edges E connecting nodes exclusively between these sets. Mathematically, no edges exist within the same set, only between them. This bipartite structure can be represented as a biadjacency matrix B, where $B_{ij} = 1$ if node i in set U is connected to node j in set V, and $B_{ij} = 0$ otherwise.

Version 1.18, September 19, 2023

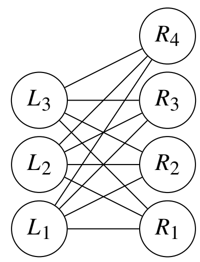

Fig. 1 A bipartite graph with nodes on the left (L_1, L_2, L_3) and nodes on the right (R_1, R_2, R_3, R_4).

2.2 Applications in Diverse Domains. The versatility of bipartite graphs is evident in their widespread applicability across various fields. In ecology, bipartite networks model interactions between species and their resources, shedding light on ecological dependencies and trophic relationships. In social sciences, they capture collaborations between individuals and organizations, offering insights into information flow and partnership dynamics.

One notable application of bipartite graphs lies in modeling electoral systems. In this context, nodes can represent voters and candidates, and edges signify the preferences or affiliations of voters towards specific candidates. This framework provides a powerful tool for understanding the dynamics of opinion formation and voter behavior, enabling insights into information dissemination and coalition building during elections.

2.3 Projection and Projection Methods. A notable feature of bipartite graphs is their potential for projection onto one of the node sets. By projecting a bipartite graph onto either set of nodes, unipartite projections are derived, enabling the study of interactions among nodes of the same type. For instance, in a co-authorship network, projecting onto the set of authors yields a unipartite network where edges denote direct collaborations between authors.

2.4 Analytical Toolbox. In order to analyze a bipartite graph one can look at some metrics:

- **Degree Distribution:**

$$\text{For nodes in set } U: \qquad P(k_U) = \frac{n_U(k_U)}{|U|}$$

$$\text{For nodes in set } V: \qquad P(k_V) = \frac{n_V(k_V)}{|V|}$$

- **Clustering Coefficient:**

$$C = \frac{\sum_{i \in U} \sum_{j \in V} \frac{e_{ij}}{k_U(i) \cdot k_V(j)}}{|U| \cdot |V|}$$

- **Assortativity:** Assortativity measures the preference of nodes to attach to others that are similar in some way. In bipartite graphs, assortativity with respect to degree quantifies the correlation between the degrees of nodes in sets U and V.

- **Modularity:** Modularity quantifies the strength of division of a network into modules or communities. In bipartite graphs, modularity is a measure of how well the nodes in sets U and V form distinct communities based on their connections.

In summary, bipartite graphs stand as a foundational tool in network science, facilitating the representation and analysis of complex relationships in a diverse range of disciplines. Their distinctive structure and versatile applications make them indispensable for modeling interactions in ecological, social, and technological systems. In the following sections, we delve into the specific application of bipartite graphs, emphasizing their role in unraveling intricate patterns and informing predictive modeling in various domains.

In summary bipartite graph constitute a powerful mathematical tool to study a plethora of complex system and quantify certain aspects of them. In following sections we will see how to apply the formalism described. Moreover we will exploit the degree metric to get insight on the result of the voter model simulations the real world problem for an election system. To get a more theoretical and rigorous description of these structures see [1].

3 Voter Model on Complex Networks

3.1 Introduction to the Voter Model. The Voter Model was first introduced by Clifford in [2] in 1973 to study the spacial conflict of species in a given space. It was formally defined in 1975 by [4], a mathematician who laid down the basis for further theoretical developments.

The Voter Model simulates the evolution of opinions or states among interacting agents in a networked system. In this model, nodes, representing individuals or agents, are connected through edges signifying their potential for influence or interaction. Over time, agents can iteratively adopt the opinions of their neighbors. This often induces interesting dynamics of consensus.

3.2 Process of the Voter Model on Graphs. It is a stochastic process used to model opinion dynamics in social networks. It can be represented as a graph $G = (V, E)$, where V is the set of nodes representing individuals and E is the set of edges denoting possible interactions. Each node v has an associated binary state σ_v, which represents its opinion.

At each iteration, a node v is chosen uniformly at random from V. Then, one of its neighbors u is selected uniformly at random from the set of neighbors $N(v)$. Node v adopts the opinion of node u, i.e., $\sigma_v = \sigma_u$.

This process continues for a specified number of iterations or until a convergence criterion is met. Over time, the opinions of nodes in the network tend to homogenize, and a consensus or polarization may emerge depending on the initial conditions and network structure.

3.3 Network Topologies and Influence Propagation. The choice of network structure profoundly impacts the dynamics of the Voter Model. Scale-free networks, characterized by a few highly connected hubs, exhibit resilience to rapid opinion shifts, promoting the coexistence of multiple opinions. On the other hand, regular lattices tend to induce consensus due to their homogeneous connectivity. Small-world networks, with a balance of local and long-range connections, offer an intriguing interplay between these extremes, fostering both diversity and cohesion in opinion formation.

3.4 Phase Transitions. As for physical states, opinion can go through phase transition. By varying network connectivity or initial conditions the system may undergo abrupt shifts from fragmented, multi-opinion states to uniform consensus or vice versa. Finding the critical points in which the consensus shift may give an insight into the stability and resilience of opinions in a network.

3.5 Extensions and Applications. Beyond its foundational formulation, the Voter Model has seen numerous extensions to account for more nuanced social interactions. Incorporating bounded confidence, external stimuli, or adaptive networks provides a more realistic portrayal of opinion dynamics in dynamic, real-world scenarios. Moreover, applications of the Voter Model extend to fields as diverse as political science, epidemiology, and marketing, where understanding the spread of opinions or behaviors is of paramount importance.

4 NBA Complex Network Analysis

The analysis is conducted using the NetworkX [3] library in Python, focusing on metrics and visualizations related to voter-player interactions.
The analysis with the voter model is to be compared to real data that can be found on the NBA official site ([5])

4.1 Choice of MVP. The NBA Most Valuable Player (MVP) is determined through a meticulous voting process by a panel of seasoned sports journalists and broadcasters. Each voter selects their top choice for the MVP award, considering factors such as individual performance, team impact, and league-wide contribution. The initial preference carries the most weight in the final tally. For this analysis we will only use this first choice for simplicity. The player with the highest count of first-place votes is honored with the MVP title, reflecting their exceptional skills, leadership, and influence during the regular season. It's important to note that voters also consider a player's sportsmanship, character, and conduct both on and off the court.

4.2 Voter Model. The model tries to approximate the complex process described above with some given hypothesis. First it is assumed that the initial opinion of each voter is based on players statistics such as points per game or efficiency (averaged over the whole yer). Then each voter is assigned a number of neighbors and changes its opinion based on the opinion of its neighbours as described in 3.2.
By simulating many times this process one can find the best parameters to approximate the system under the assumptions above.

4.3 Analysis of the NBA data. For simplicity the only data used was that of the first choice for each voter, but in general this analysis can be extended to the more complex weighted voting process. Using this subset of the data in [5] we have plotted the bipartite graph:

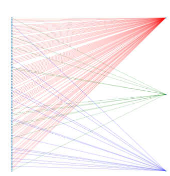

Fig. 2 Graph rapresentation of ballots from ground truth. RED → Embid, GREEN → Giannis, BLUE → Jokic

then through the degree for player nodes one can visualize how many votes each player recieved. It is worth noting that the degree for each voter is precisely one since each voter casts only one vote

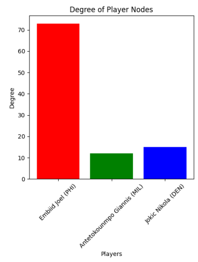

Fig. 3 Degree for players with colors associated to each player

4.4 Voter Model. Firstly we must define the assumptions of the model. Each voter initializes its opinion with probability to choose a player proportional to a scalar called **value**. This scalar contains information about the performance of the player during the whole season. This parameter can be defined in various ways, for this analysis the formula chosen is

$$value = (efficiency + \frac{assists + points}{10}) \cdot minutes$$

. Once initialized each voter chooses 2 neighbours which are the ones nearest to him in the list of voters and randomly changes its opinion as described in 3.2.

In order to obtain more reliable result from the simulation, all the results reported are taken from the average values across multiple simulations (100).

4.4.1 Parameters. Many parameters can be changed to obtain different results, but concerning exclusively the graph aspect the main one is the graph topology itself: in this case the choice of neighbours. Changing neighbours after each iteration brings the process to the degenerate case in which it converges extremely fast (less than 5 iterations) to consensus even for few neighbours.

Another relevant parameter is that of the number of the graph's topology: the number of neighbours. It might be interesting to simulate the process with all possible choiches of neighbours in the voter nodes, but due to lack of computational power only a handful of values have been tried. The results shown have a fixed number of neighbours equal to four.

5 Results

For a large number of iterations, each voting process tends to reach consensus as shown for a simulation with number of iterations = 500. The state which most intrests us is the one in which consensus in yet to be reached

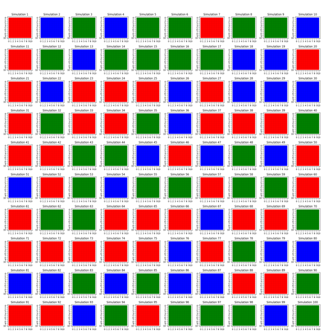

Fig. 4 Grid rapresentation of ballots in the limit case. RED → Embid, GREEN → Giannis, BLUE → Jokic

The state which most interests us is the one in which consensus in yet to be reached, so for a smaller number of iterations (>10, <20). Below is shown the result of a simulation with number of iteration = 17

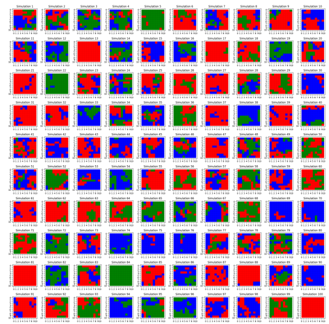

Fig. 5 Grid rapresentation of ballots in a simulation with number of iterations = 17. RED → Embid, GREEN → Giannis, BLUE → Jokic

then through the average degree for player nodes one can visualize how many votes each player recieved on average. Clearly also for this case the degree for each voter is precisely one since each voter casts only one vote

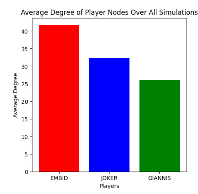

Fig. 6 Average degree for players with colors associated to each player

6 Conclusions

The application of the voter model in these case has generated results that can be compared with the ground truth:

Table 1 Comparison between model and ground truth

Player	Simulation Degree	Real Degree
Embid	41.63	73
Jokic	32.35	15
Giannis	26.02	12

This discrepancy can be due to many different factors: different statistical analysis for the value formula by the voters, dynamic topology of the network, personal bias of the voters and many others unknown ones.

Therefore, despite the fact that the results shown somewhat resemble the ground truth, the conclusion of this analysis is that this formulation of the voter model is unfit to make accurate prediction on the ballot's process for MVP in the NBA

References

[1] Béla Bollobás. *Modern graph theory*, volume 184. Springer Science & Business Media, 1998.

[2] Peter Clifford and Aidan Sudbury. A model for spatial conflict. *Biometrika*, 60(3):581–588, 1973.

[3] Aric Hagberg, Pieter Swart, and Daniel S Chult. *Exploring Network Structure, Dynamics, and Function using NetworkX*, 2008.

[4] Richard A Holley and Thomas M Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *The annals of probability*, pages 643–663, 1975.

[5] NBA. https://pr.nba.com/voting-results-2022-23-nba-regular-season-awards, 2023.

List of Figures

List of Tables