Analyzing the network-ecosystem of competitive online video-games.

Term Paper: 'Network Science of Socio-Economic Systems' by Robin Aytac

April 2025

1 Abstract

This report presents an analysis of a digital ecosystem in the form of a character social-network from a Pokemon Online Battle Simulator. We construct a new dataset and analyse the topological properties of the resulting graph. Furthermore we analyse the relationship between topological properties, popularity and in-game effectiveness. We make the argument that this digital ecosystem shows similar behaviour as real-life ones with a trade-off between specialization and versatility governing ecological fitness.

2 Introduction

A key idea in video games is the idea of a 'meta-game' (Most Effective Tactic Available), meaning a strategy that yields optimal results when compared to all others. A meta-game is usually not formally defined but an emergent property of a social system. Players develop intuition and share information which leads to a centralisation of the decision space towards these efficient strategies without any centralised effort. Player vs. Player (PVP) games are particularly interesting as we can observe strategies and counter-strategies. Pokemon is a special case of a PVP game as the game can be defined in two distinct phases. The core gameplay consists of using your team of 6 characters to beat your opponents team using good decision-making and taking strategic risks. The more complex phase is however the team-building. Each player chooses and customizes 6 characters from a finite set of options. The decision-space for this is too vast to be sensibly explored. In the current standard format the selection of team-members alone is in the realm of greater than $\binom{800}{6}$ or approx. 3×10^{15}. Despite this theoretical intractability we can observe the development of a meta. As every battle can be considered a random experiment in this decision space with a clear win-lose outcome, players learn what good and bad strategies look like over time. This results in certain combinations being picked much more often than others. We will use randomly sampled data from the biggest Pokemon

Figure 1: Biggest Component of post-validation Graph, colored by viability - rank.

Battle Simulator PokemonShowdown [1] to explore team-building as a network of co-occurrences. We will attempt to identify patterns in the topology of the graph that yield insights on the game and investigate whether these topological patterns translate to real world results.

2.1 Dataset Construction

The Dataset was collected by scraping online games from the PokemonShowdown replay API[2]. For that purpose I scraped approx. 50 games a day for the entire month of January 2025, for a total of 1683 games, yielding 3366 teams using the replay text. The replay text is processed by using regular expressions to extract each players team of usually 6 members, the winner of the game and the rating average between the players. For each team (read: 2 per game) we then update co-occurrence statistics. Meaning for all 15 combinations of a 6 member team, an edge is either created or updated by incrementing the weight by 1. This leads to a weighted, undirected graph of 517 nodes and 12600 edges, with the edges weighted according the sample co-occurrence of the edge.

There are some special considerations with regards to our dataset: If we view team-building as sampling 6 units from a limited node-space without replacement it is easy to make the argument that a graph constructed in such a manner will be close to fully-connected as we increase the sample size. As we do not have access to every game but are instead building a limited sample many of these

[1] https://pokemonshowdown.com
[2] https://replay.pokemonshowdown.com

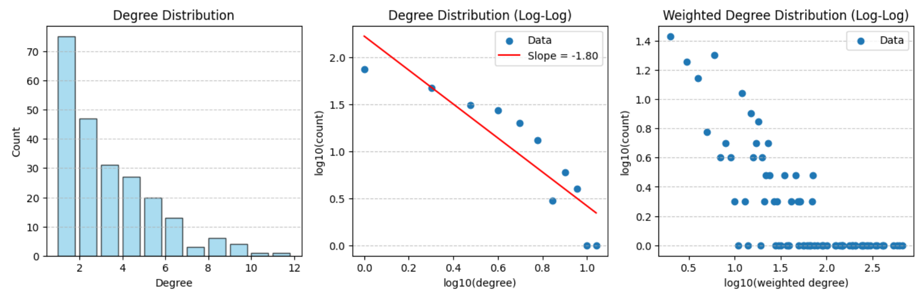

Figure 2: Degree Distributions

edges and statistics will be based on spurious information. To distinguish signal from noise I considered 2 options: Ad-hoc filtering or model-based filtering. For ad-hoc filtering we consider the average rating of each battle and only include games above the rating threshold of 1100 ELO. This will limit the sample size to only "high-quality" information as we can assume that higher-ranked games are more informative. This also very effective in filtering out one-off cases as the distribution of unique Pokemon is tighter than in the unrestricted dataset. Model-based filtering on the other hand offers a more rigorous, but less straightforward option. Instead of manually projecting the graph from the team data, use the Bipartite Configuration Model (BiCM). [6] We generate a null-model based on our data by randomly reshuffling edges under the constraint of preserving the degree of each node. This null model allows us to identify statistically significant edges ($\alpha = 0.01$) and discard others, consequently filtering all nodes that have no robust edges at all. This leaves us with 228 nodes and 337 edges. This process disconnects the graph and we go from one connected component to 20. These are however mostly negligible: The biggest connected component includes 171 nodes (75%) and with that the vast majority of interesting data. The next largest is of size 5 (2.1%). When considering the data source we can consider these small components to be artifacts of the sampling procedure. They mostly consist of Pokemon generally not considered to be competitively viable.

3 Main Analysis

3.1 Descriptive Graph Overview

Figure 1 visualizes the largest component of our graph. The nodes are coloured according to their "viability-rank" a rough metric based on ongoing community votes that sorts Pokemon into hierarchical tiers [3]. We start by observing the degree distribution of the graph. We see it roughly follows a power-law distribution which is in line with many other real world graphs [1]. Figure 2 shows that

[3]https://www.smogon.com/forums/threads/sv-ou-indigo-disk-viability-ranking-thread-update-on-post-927.3734134/

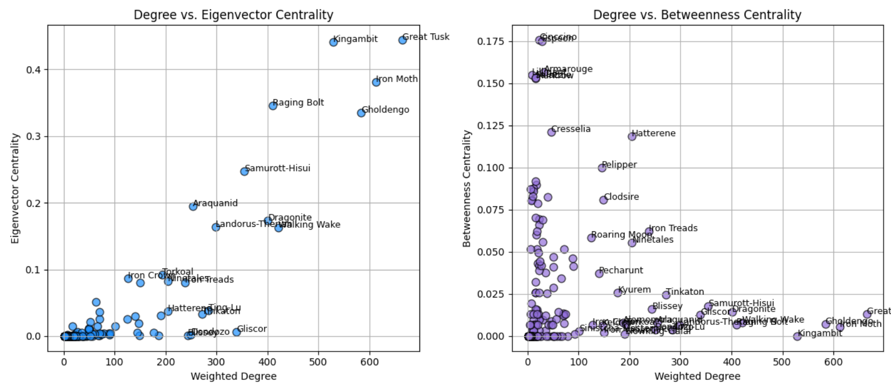

Figure 3: Centrality of Nodes

the log-log scale best fit of the degree distribution is best described by a slope of - 1.8 which allows us to estimate the shape parameter of it's distribution at around $\gamma = 1.8$ [2]. Substantively we can infer from this that there is a small number of strong hubs in our graph. They dominate the network structure and are surrounded by a large number of weakly connected nodes. We compute a weighted assortativity coefficient of 0.5. This strong positive coefficient informs us about the nature of the connections in the graph as it implies that nodes with similar degrees cluster together. From this we can substantively conclude the following: We observe a highly centralized game where a handful of elite Pokemon are connected to each other and from which tree-like substructures point away from the center. These subtrees may represent specific strategies that are essentially glued together by key Pokemon. This perspective makes intuitive sense when considering the visualization of Figure 1.

Key players can fulfil two roles from a graph perspective. They can act as structurally redundant nodes in dense regions, or act as bridges connecting different dense regions. Substantively this means, these Pokemon are either just very good in many constellations or they open up specific strategies. We will compare three metrics: Weighted degree, eigenvector centrality and betweenness centrality in order to get an overview.

We immediately see that these two centrality measures are not equivalent. In the eigenvector-centrality plot we observe a linear relationship between weighted degree and centrality. We see high-degree nodes are the most central ones as they fill up the dense center of the graph. These are very strong all-rounders that can find a place on many different teams, and are favoured by many players. In the betweenness-centrality plot we observe a systematic but non-linear relationship. Nodes with low degree show high betweenness, meaning these are rarely used but important in opening up specific areas of the graph allowing for the use of niche strategies in which they are non-replacable. On the other hand we can observe a number of high-degree low betweenness elements that are structurally redundant as there are many team configurations in which one of these nodes

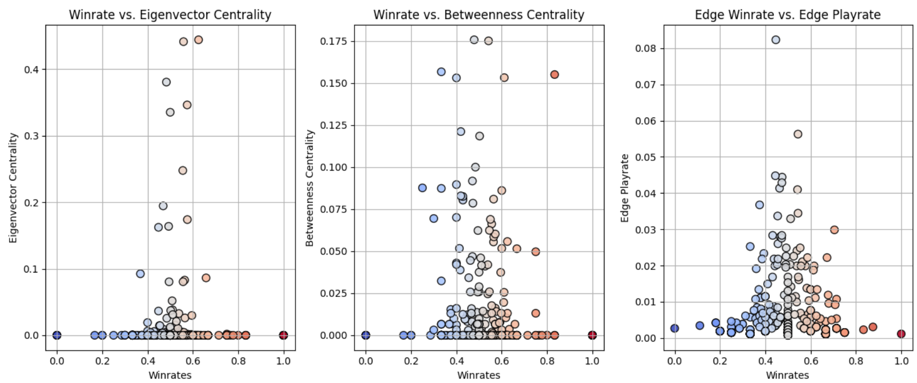

Figure 4: Winrate vs a) Centrality b) Betweenness c) Edge playrate

is replaced by another. This strengthens the above point about a cluster of highly-connected nodes dominating the graph structure.

3.2 Win-rate Analysis

Figures 4a) and 4b) present node winrates. We can observe a nonlinear funnel-like shape looking roughly gaussian. We observe that Pokemon occupying central positions in the network, display middling performance. While middling win-rate Pokemon have high centrality. Figure 4c) shows that this pattern extents to edges - so pairs of two Pokemon as well. Again we observe a large number of middling but popular pairings. These can be considered broadly applicable, consistent duos. Meanwhile, less commonly used combinations are either displaying underrated synergies not picked up by the majority of the playerbase, or ineffective combinations, often called "traps".

4 Discussion

An interesting pattern emerges here. We see that the central nodes in our graph appear to be the most highly-ranked ones by the community. But we also see that the most central nodes are not necessarily the ones that win the most. We can think about the network of Pokemon as a digital ecosystem for a nature inspired analogy [3]. We can divide the hub-and-spokes topology of our network into central and peripheric Pokemon. The most important players in this ecosystem are staples that perform consistently with a win-rate hovering around 50%. They are "fit" in terms of this ecosystem by offering a stable basis for others to build around. But as we see by their middling win-rates they are so common that their individual impact is quite hard to assesss. Based on this we can assume that the key nodes identified before are what can be considered a baseline when building a team. They are used a lot and can serve as easy to use staples that stabilize any team composition. On the other hand, Pokemon with very high win-rates are not central at all, implying they are only present in highly specialized teams. Niche or combo-dependent Pokemon can yield excellent results but they can also perform much worse than their more robust peers. The provided analysis provides a great opportunity for identifying "overperformers" that can yield strong results even though they are considered to be on the periphery of the metagame. We observe here a tension between versatility and specialization [5]. In our ecosystem, the consistently high use of versatile elements indicates that while specialization can drive exceptional performance in specific contexts, it is the balanced integration of a range of competencies that enhances overall resilience and innovation. Such a dynamic is observed not only in digital networks but also across biological, social, and economic systems, where the tension between narrow specialization and generalist adaptability is key to evolutionary success and system sustainability [7][4].

References

[1] Albert-László Barabási and Márton Pósfai. *Network science*. Cambridge: Cambridge University Press, 2016. ISBN: 9781107076266 1107076269. URL: `http://barabasi.com/networksciencebook/`.

[2] Béla Bollobás* and Oliver Riordan. "The diameter of a scale-free random graph". In: *Combinatorica* 24.1 (2004), pp. 5–34.

[3] Matthew D Holland and Alan Hastings. "Strong effect of dispersal network structure on ecological dynamics". In: *Nature* 456.7223 (2008), pp. 792–794.

[4] Pietro Landi et al. "Complexity and stability of ecological networks: a review of the theory". In: *Population ecology* 60.4 (2018), pp. 319–345.

[5] Pratha Sah et al. "Exploring community structure in biological networks with random graphs". In: *BMC bioinformatics* 15 (2014), pp. 1–14.

[6] Fabio Saracco et al. "Randomizing bipartite networks: the case of the World Trade Web". In: *Scientific Reports* 5.1 (June 2015). ISSN: 2045-2322. DOI: `10.1038/srep10595`. URL: `http://dx.doi.org/10.1038/srep10595`.

[7] Andrea Tacchella et al. "A new metrics for countries' fitness and products' complexity". In: *Scientific reports* 2.1 (2012), p. 723.

5 Appendix

5.1 Code and Data

Code and Data can be found on GitHub `https://github.com/AytacRB/Pokeproject`

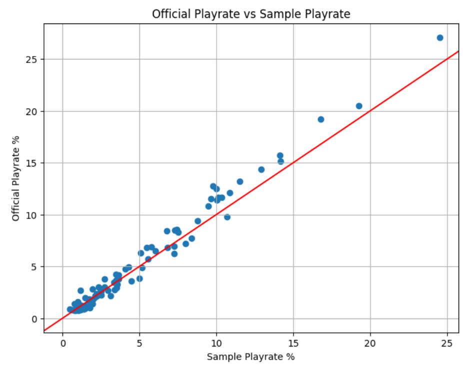

Figure 5: Scatterplot of official playrate and our sample data playrates

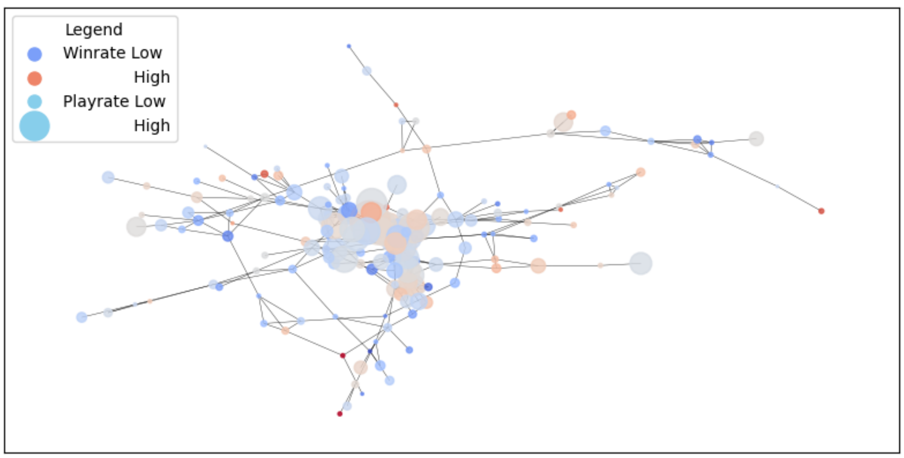

Figure 6: Node-properties displayed on graph

5.2 Validation with official data

While there is no official information for winrates or co-occurance, Pokemon-Showdown releases monthly usage statistics [4]. We can compare our sampled data with the official data in Figure 5 to see that we can assume high data-quality due to a very high correlation with the official statistics.

5.3 Node-statistics on graph

For better understanding of the distribution of node-level statistics across the graph please inspect Figure 6.

[4]January 2025: https://www.smogon.com/stats/2025-01/gen9ou-0.txt