**Do the existing categories in this abridged Wikipedia reflect optimal shortest paths and/or the path effectively taken by players?**

**Optimal paths**

We first consider the optimal shortest paths within the Wikipedia link network. We compare the network properties of the linking relations over all articles and of the network of articles within a single category.

We see that some categories are larger (more nodes) or have more edges than the expected average (total divided by 15).

We see that, within each category, the average degree is lower (which is reasonable since intercategory-links are excluded). On the other hand, the density and clustering coefficients are mostly higher within each category. This suggests that articles within the same category are more interconnected, forming closer networks compared to the overall Wikipedia link network.

We then look into the optimal paths. The average shortest-path length in the overall network is 3.2. Within each category, the average SPL ranges from 3.03 to 3.29, with a mean of 3.18 and a median of 3.16. The SPLs in each category are analyzed using ANOVA with a statistically significant result, i.e., the variability between category means is larger than the variability of the observations within the categories.

These results indicate that the existing categorization does form reasonable clusters that group articles which are more closely related, although without an alternative comparison, it is hard to conclude how “good” the clustering is.

**Player paths**

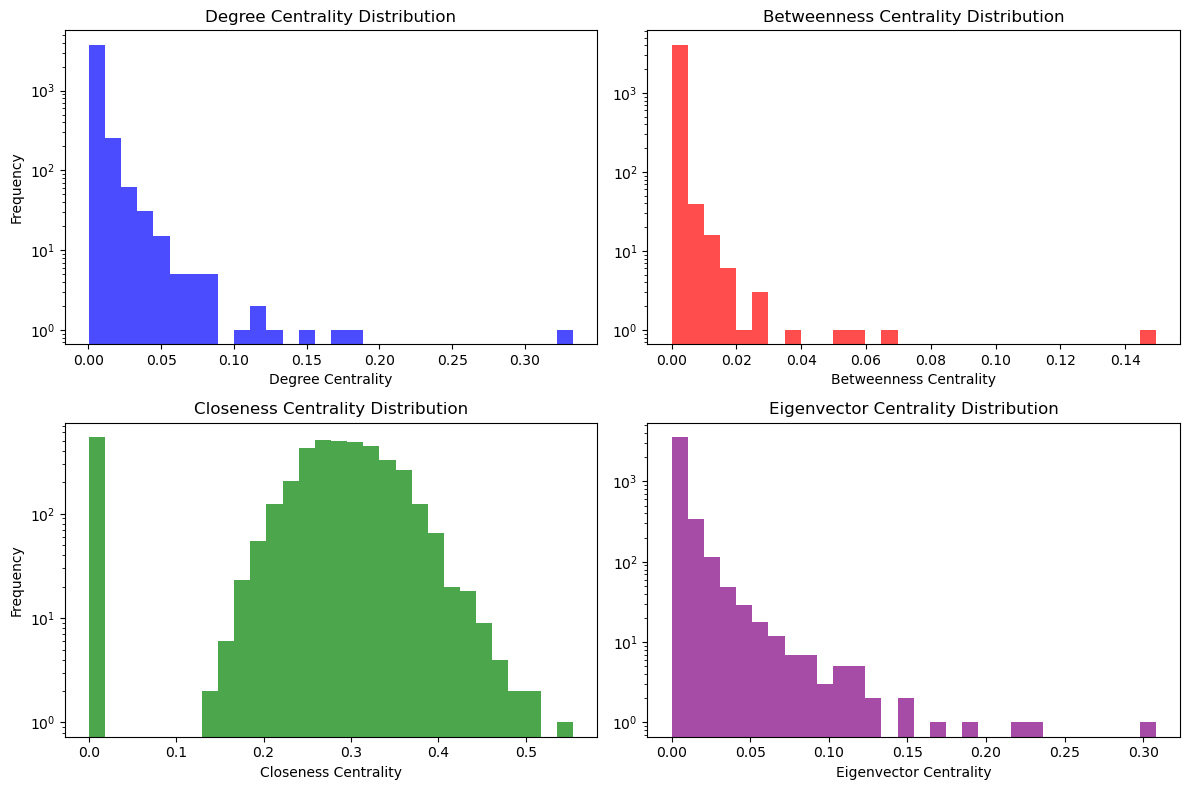
To investigate whether the existing categories reflect the paths taken by the players, we can look into how the categories affect player’s performance. If the categories reflect the paths taken by the players, we can hypothesize that the paths for articles in the same category should be shorter and the player may perform better (e.g., taking less time).

If the starting and target articles are in the same category, the chance of finishing a path increases by 7.3% (statistically significant) assuming a linear model. Among the finished paths, being in the same category results in on average a 31-second reduction in game completion time (significant), again assuming a linear model. The category also affects the difficulty rating given with an estimated average effect of -0.19 (less difficult) for having the same category of starting and target articles.

Given these statistically significant effects, we can conclude that the existing categorization does reflect the player’s strategy and view of difficulty.

**How do different network centrality and clustering measures affect player paths and can it give us insight into the way humans perceive and utilize hubs in an information network?**

To answer this question, we investigate the nodes most utilized by the players and compare them to the “presumed” hubs using different centrality and clustering measures.

The different centrality measures for paths taken by the players are plotted below. They are similar to the distribution in general.

We then compare the distribution of different centrality measures for most frequented 100 articles in player paths. In all cases, the mean and median centrality values are higher for these hubs. Betweenness and eigenvector centrality measures have a much smaller value range, and the difference for hub articles are less easily seen. Closeness centrality has a “nicer” distribution compared to others with heavy tails and may be more suitable to qualify hubs.

Additionally, the top nodes on average have more categories (1.28 per article instead of 1.13).

