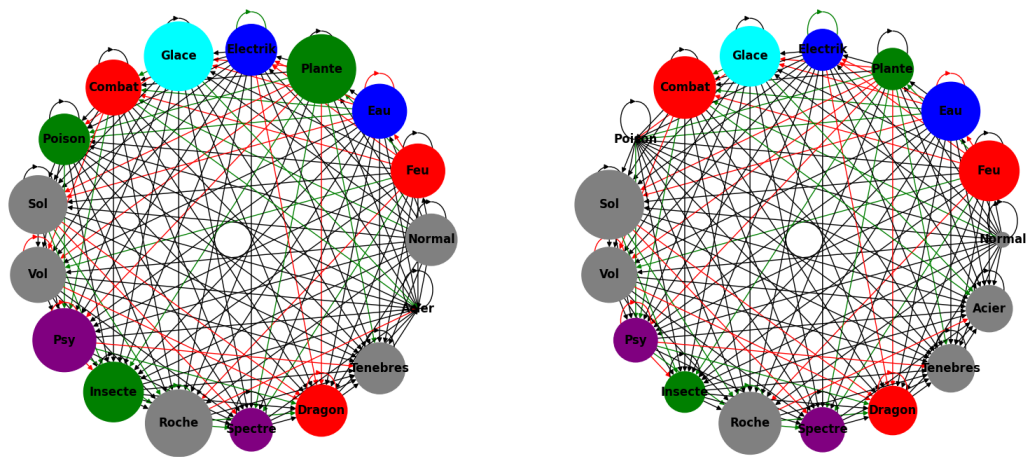


PageRank

Vincent Foriel

December 10, 2021



Contents

1	Introduction	3
2	The PageRank vector	3
3	The CheiRank vector	6
4	PageRank : Pokémon, link them all!	7
5	Conclusion	8

1 Introduction

The objective of this practical work is to study the PageRank algorithm by applying it in simple situations.

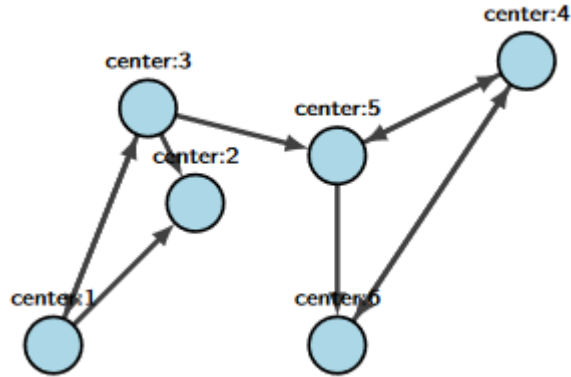


Figure 1: A simple oriented graph

2 The PageRank vector

We consider the graph represented in Fig 1. In our exercise, each node represent a website, and each link represent a button or an hypertext link that redirect to another website. The objective is to determine the probability of being in each nodes after a huge number of steps (one step represent a redirection on another website)

For a random surfer that start his journey on a random node and with a dumping factor at 0.85, he have these chances to finish at each node:

Node 1 : 0.0517050169447207
Node 2 : 0.0736797339206392
Node 3 : 0.0574127283822214
Node 4 : 0.3487031406768242
Node 5 : 0.1999036807831220
Node 6 : 0.2685956992924724

This stability is reached after 23 steps.

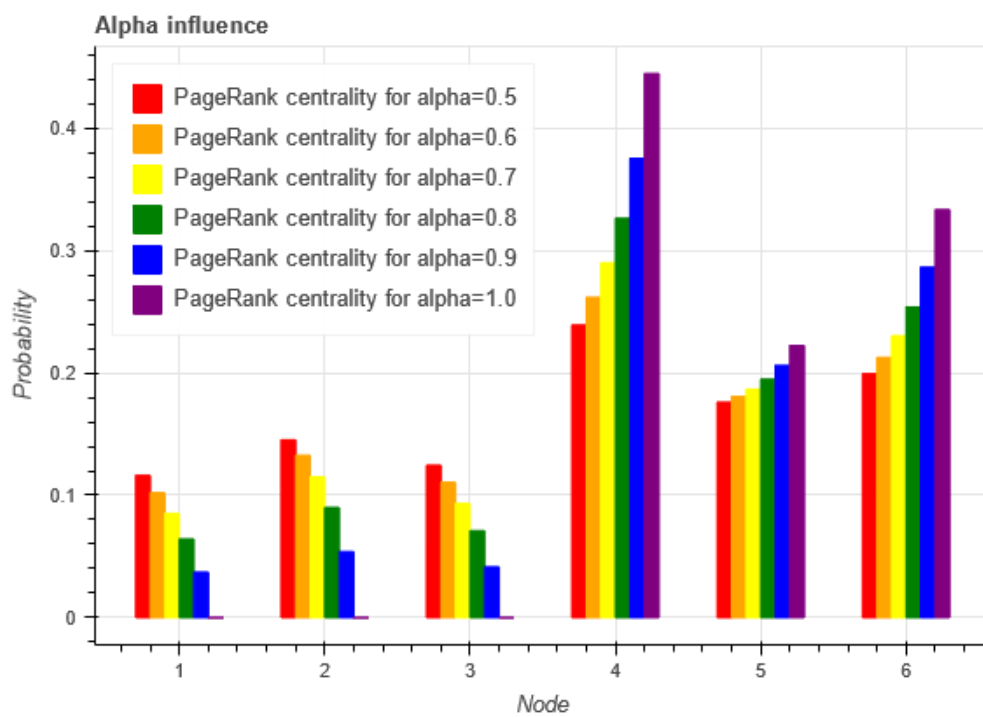


Figure 2: Influence of the dumping factor

Now, if we change the dumping factor and makes it increase from 0.5 to 1, we can see that the probability change i.e. Fig 2. The greater the dumping factor is, the more chances the surfer has to leave the left part of the graph and stay on the right part.

In comparison, degree centrality gives us:

Node 1 : 0.15
Node 2 : 0.10
Node 3 : 0.20
Node 4 : 0.20
Node 5 : 0.20
Node 6 : 0.15

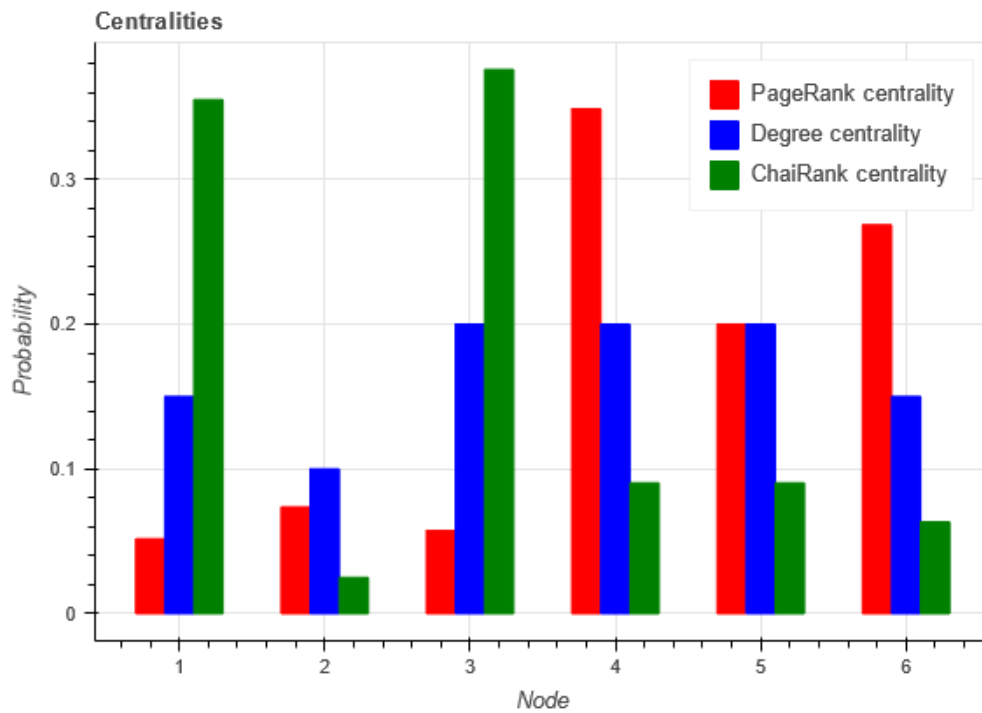


Figure 3: Centralities for the different methods

3 The CheiRank vector

For CheiRank, with a dumping factor of 0.85, we got the following probabilities:

Node 1 : 0.3550956514060691
Node 2 : 0.0250000000000000
Node 3 : 0.3758472962936703
Node 4 : 0.0903328050722020
Node 5 : 0.0903328050722020
Node 6 : 0.0633914421558570

This stability is reached after 68 steps.

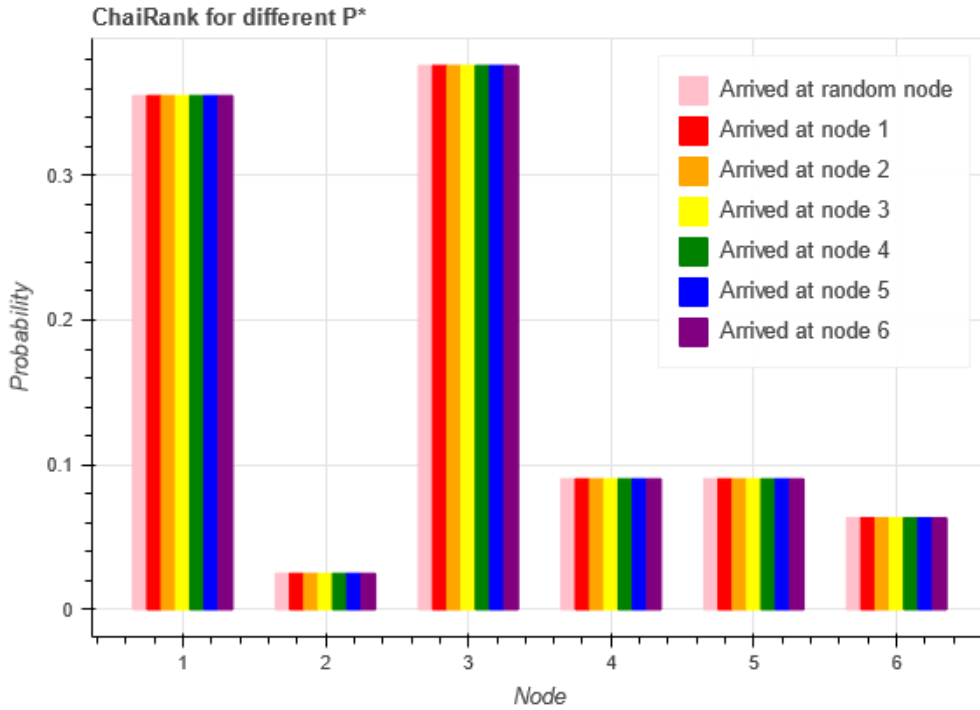


Figure 4: Best referrers depending of the arrival website

We can observe the difference between the 3 methods in the Fig 3. As we expected, the CheiRank vector seem to give more importance to the node

that have the most outgoing links. My interpretation is that this vector can allow to create a referencing of biggest referrers.

By changing the arrival website, we can see on the Fig 4 that the probabilities of the referring website are exactly the same. As the Fig 5 show, the CheiRank vector tends quickly to it's final state. We can draw the same graph with different starting node, we always see the same convergence values.

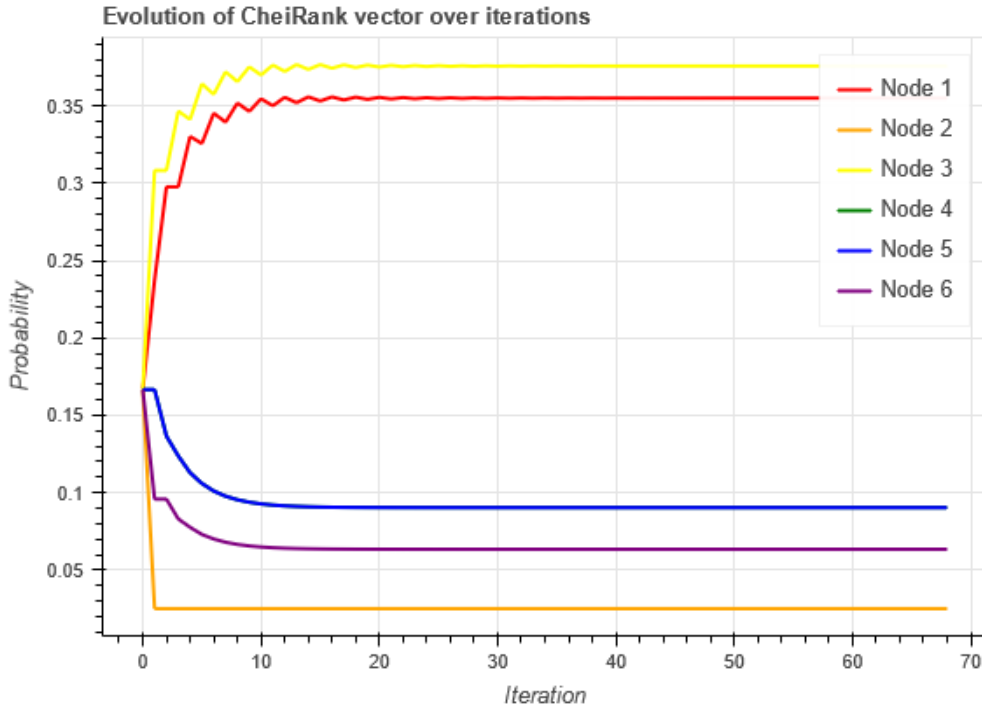


Figure 5: Evolution of the CheiRank vector

4 PageRank : Pokémon, link them all!

The adjacency matrix represents basically the links and their respective "strength" between the different network nodes. The stochastic matrix is more interested in the number of possible ways to leave/arrive from a given node.

5 Conclusion

For the first graph, the objective of the PageRank (resp. CheiRank) seems to be the ranking of the most referred (resp. referrer) websites. However, when we use them for a more complexe graph like the pokemon one, the interpretation becomes less evident.

At first sight, by looking the lines of the matrix, I noticed that the Pokemon that is the best ranked in the CheiRank vector seems to correspond to a Pokemon that have one of the greatest number of critical damages with not many unefficients attacks. I computed the ratio of attack and defense and I put the data in a table (in the data/analysis.ods file) to see the link between the rankings and these ratio. We can see that the CheiRank seems to correspond to the types that have the greatest damage ratio, and the PageRank seems to correspond to types that have the best defense ratio. However, there is some little exceptions that I can't explain with this hypothesis.

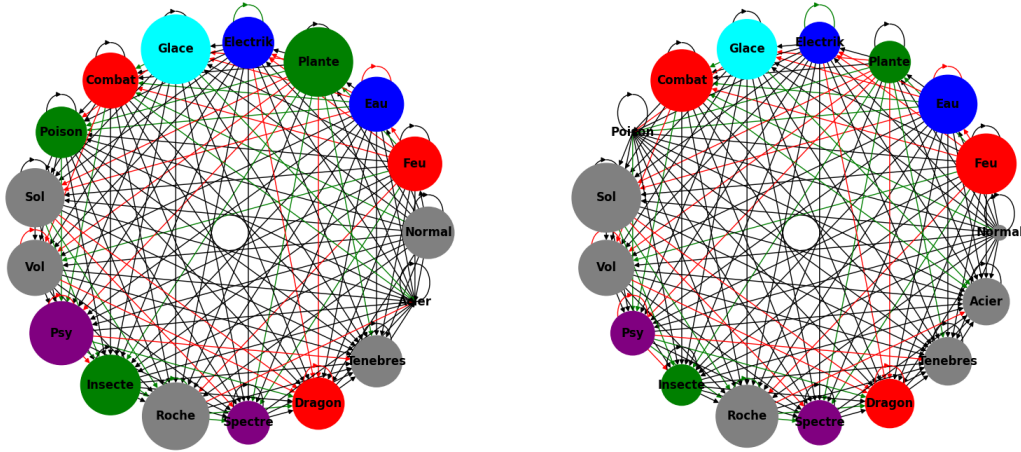


Figure 6: Pokemon interactions and ranking with PageRank (left) and CheiRank (right)

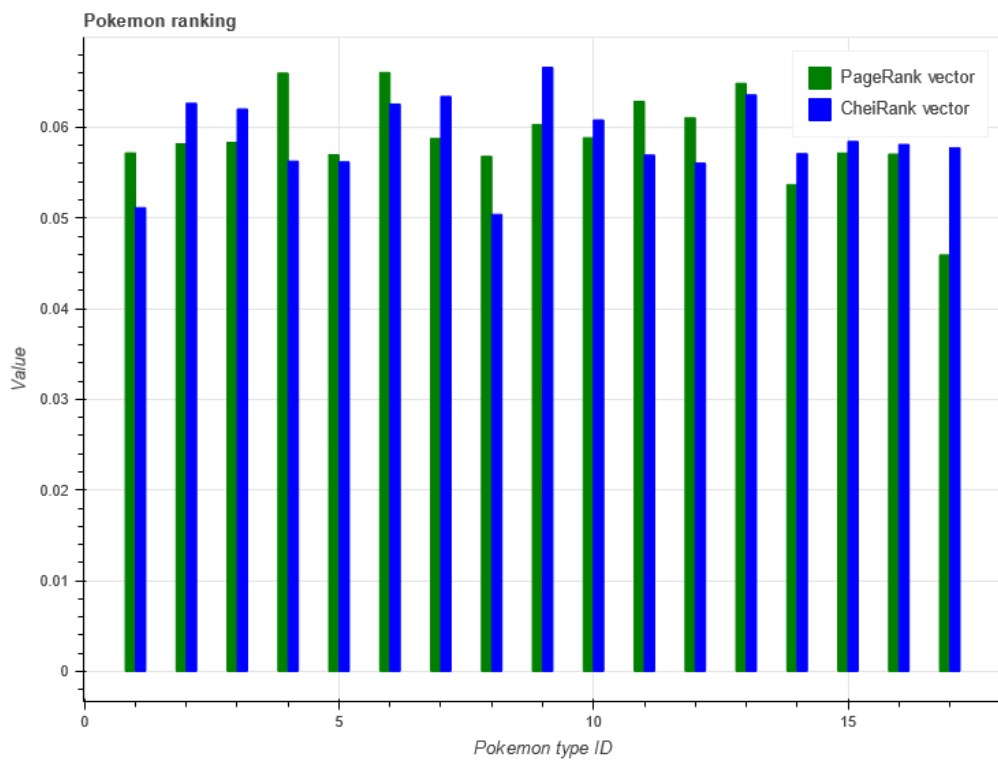


Figure 7: Pokemon ranking