Power low distribution in dynamic and static network

Yishai Coorsh

The project examines whether the power law distribution exists in static networks and not in dynamic networks in an empirical way. To test this, we used Clauset methodology for an accurate examination of the distribution.

The methodology includes three stages:

1. Estimate the parameters xmin and α of the power-law model.
2. Calculate the goodness-of-fit between the data and the power. If the resulting p-value is greater than 0.1 the power law is a plausible hypothesis for the data, otherwise it is rejected.
3. Compare the power law with alternative hypotheses via a likelihood ratio test.
   For each alternative, if the calculated likelihood ratio is significantly different from zero, then its sign indicates whether the alternative is favored over the power-law model or not.

We examined static and dynamic data from three different data:

1. **Edits Wikipedia article**
   The source of data its from dumps Wikimedia[1] Correctly updated to 1.8.2021. The data was processed to localize the page articles according to Wikipedia definition[2].
2. **Stack Overflow temporal network[3]**
   Count per user answered question, commented question and commented answer. for each user the date of the last activity was identified
3. **Arxiv High Energy Physics paper citation network[4]**
   citation graph is from the e-print arXiv and covers all the citations within a dataset of 27,770 papers with 352,807 edges.

Result

	Edits Wikipedia article		Stack Overflow temporal network		Arxiv High Energy Physics paper citation network	
	dynamic	static	dynamic	static	dynamic	static
α	2.14	3.5	2.27	2.83	3.47	3.5
Xmin	122	206	380	946	94	41
N	6,314,051	2,164,548	2,226,242	1,299,062	24,833	17,616
p-value (goodness-of-fit between)	0.00	0.12	0.00[5]	0.9992	0.7432	0.3232
D	0.022	0. 00761583	0.012	0.0069	0.0237016	0.0299562

		power law	lognormal		exponential		Stretched exponential		Truncated power_law		support for power law
		P	LR	p	LR	p	LR	p	LR	p	
Edits Wikipedia article	dynamic	0.00	-	-	-	-	-	-	-	-	None
	static	**0.12**	-1.37	0.17	11.3	**0.00**	2.26	**0.02**	-2.5	**0.02**	With cut-off
Stack Overflow temporal network	dynamic	0.00	-	-	-	-	-	-	-	-	None
	static	**0.9992**	-0.34	0.73	**7.32**	**0.00**	1.32	0.18	-0.65	0.25	Moderate
Arxiv High Energy Physics paper citation network	dynamic	**0.7432**	-0.88	0.37	**2.67**	**0.00**	-0.65	0.5	-1.56	0.07	Moderate
	static	**0.3232**	-1.54	0.12	1.3	0.19	-1.55	0.11	-4.49	**0.00**	With cut-off

"good" indicates that the power law is a good fit and that none of the alternatives considered is plausible.
"moderate" indicates that the power law is a good fit but that there are ther plausible alternatives as well.
"with cut-off," meaning that the power law with exponential cutoff is clearly favored over the pure power law.
"None" indicates data sets that are probably not power-law distribute

Conclusion

The tests we did show that there is a difference between a dynamic and static network on the same data, these differences show that dynamic networks do not comply with the power law, while static networks do, even if not perfectly. In the Wikipedia data we clearly see that the dynamic network does not obey the power law, while the static network of the data clearly does, although not perfectly.

Examining the data from Stack Overflow, we see that the dynamic network does not comply with the power law at all, while the static network does. The static network test was given a "moderate" value because it did not rule out all alternative extinctions. But for our examination this result is satisfactory.

As for the data from the Arxiv High Energy Physics paper citation network, the results are divided. The dynamic network was adapted to the power law and so was the dynamic network. A possible explanation for this result is that the network is much smaller compared to the other discrete data, so the change between the dynamic and static part is not large.

Another and more appropriate explanation is that the data were collected over 10 years, which is a relatively short time, which resulted in an extremely small amount of static articles with a high amount of citations.

---

[1] Dumps wikimediai

[2] Manual:Article count, 2022

[3] Stack Overflow temporal network data source

[4] Arxiv High Energy Physics paper citation network data source

[5] Because of running time constraints. A test was performed on 100 samples for initial testing.