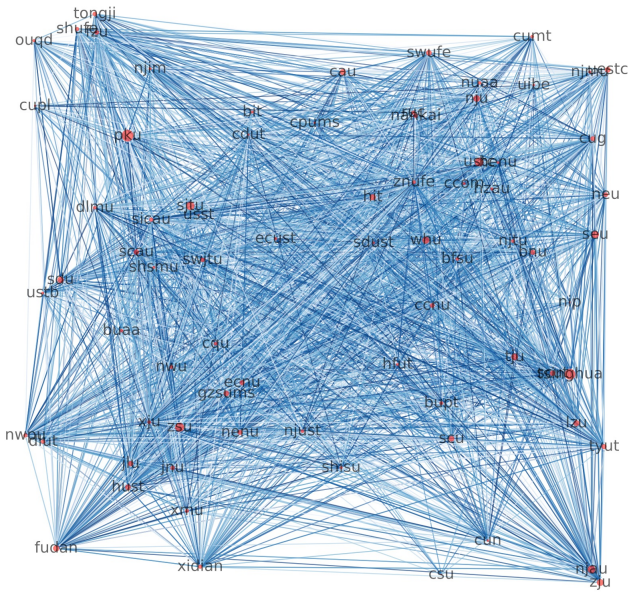


# CSIC 5011 Final-Project 1: PageRank and Primary Eigenvectors

Chutian Huang<sup>1</sup>, Mengyue Zha<sup>2</sup> {chuangat,mzha}@ust.hk  
Department of Mathematics, HKUST

## 1. Introduction

We apply Google's PageRank and HITS on Chinese University Websites. Hyperparameter alpha is tuned with a deep investigation on its effects. We also compare the ranking results against the university's research ranking by Spearman's rho and Kendall's tau. Finally, the links between universities are visualized.



## 2. univ\_cn Dataset

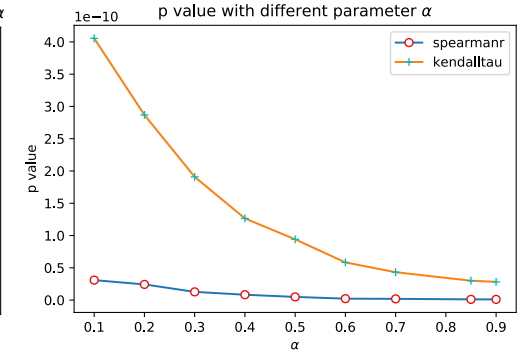
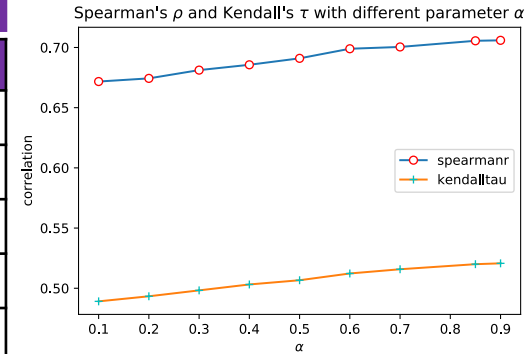
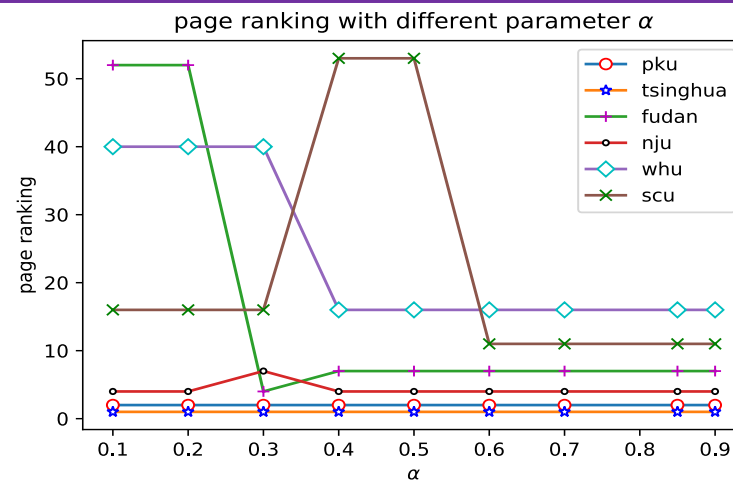
The following dataset contains Chinese (mainland) University Weblink during 12/2001-1/2002. rank\_cn is the research ranking of universities in that year, univ\_cn contains the webpages of universities, and W\_cn is the link matrix whose (i, j)-th element gives the number of links from university i to j. There are totally 76 university recorded in this dataset.

## 3. Top Universities With 4 Rankings

Ranking	Research	PageRank	HITS authority	HITS hub
1	pku	tsinghua	tsinghua	pku
2	tsinghua	pku	pku	ustc
3	fudan	sjtu	uestc	zsu
4	nju	nju	sjtu	sjtu
5	zju	uestc	nju	zju

$\rho$	$\tau$	Research	PageRank	HITS auth	HITS hub
Research			0.5200	0.5719	0.3782
PageRank	0.7056			0.4063	0.2716
HITS auth	0.7505	0.5719			0.3165
HITS hub	0.5396	0.4043	0.4413		

## 4. Alpha Tuning



## 5. Conclusion

We found that some top universities are academic centers such as Peking university, Tsinghua University, Nanjing University and Zhejiang University who appear on all. Top 5 list ranked by four methods including those aiming at finding the authorities and hubs. We further compared the correlation between ranking results by Spearman's rho and Kendall's tau. We see that when using rho, ranking results tend to have larger correlation. P-value is also involved in our comparison.

Then, we make experiments on different alpha and observe the change of correlation and p value with different alpha. We evaluate by both Spearman's rho and Kendall's tau with the research rank. We can see that as alpha increases, both correlation factors increase, and the p value decrease, which means the page rank tend to the research rank more and more closely.

Finally, we observe the variation of page ranking of six representative university with the change of hyperparameter alpha. We find that for absolute top universities like Tsinghua University and Peking University, the ranking has little variance while for Wuhan University and Sichuan University, the variances are quite large. So choosing alpha is important for most universities.

## 6. References

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The PageRank citation ranking: Bringing order to the web*. Stanford InfoLab.

## 7. Contribution

- HUANG Chutian: Coding and Theoretical Support
- ZHA Mengyue: Theoretical Support and Making Poster