Effects On PageRank Scoring Under Power-Law Step Distributions

Thunradee Tangsupakij, Geeratigan Arsanathong

Abstract— **In this research monograph, we aim to study the effect of altering the step distribution implicit in the Personalized PageRank algorithm. After modifying the step distribution within Personalized PageRank to follow Zipf's Law, we test assumptions of convergence to a stationary distribution. Furthermore, we design two distinct measures to analyze the differences between the modified algorithm's scoring in comparison the classical algorithm. Our analysis is conducted over three real-world datasets, where we find mixed results. Finally, we illustrate an interesting particular case and provide exciting possible new research directions.**

I. INTRODUCTION

Invented in the late 1990's, PageRank[1][2] emerged as a dominant method in the ranking of webpages. Historically known as Eigenvector Centrality, PageRank is named after one of its inventors Lawrence Page who later became one of the founders of the tech giant Google. Since it's inception, PageRank has been extended in a number of ways in efforts to either make a recommender system more accurate [3][4][5], or to speed up the algorithm's computation [6][7].

The mathematical theory behind ranking systems begins by viewing the system of interest as a network, where the nodes of the network may be viewed as items of the system such as webpages, social network profile pages or cities. Accompanying these nodes is the corresponding edges that link the nodes to each other: from the above examples, these can be links on webpages, friends connected to a social network profile or roads connecting each city. With a mathematical structure established for the system of interest, the goal is to create an efficient searching method in order to identify a node or a set of nodes of interest to a possible user's query search. Search algorithms can be typically boiled down to two steps; in the first step, the items in the query are searched over all nodes in the network to either identify nodes containing the exact terms in the query, or to find nodes with some relation to the query. This can be achieved in a number of ways using basic text processing. These methods include, but are not limited to using basic look-ups into a filing system or more advanced methods such as vector space methods [8]. Despite the deceiving simplicity the problem presents in the first step, often times the network being searched over consists of a massive number of nodes. For example in the case of webpages, there exists over 1.2 trillion webpages as of January 2020 [9]. In terms of social networks, the popular social network Facebook reportedly has over 2.6 billion active users (nodes) in the first quarter of 2020 [10], where the average number of friends (edges) each profile has contains is 338 [11]. The sheer size of these networks motivates the need for a ranking system in query searching, thus driving the purpose behind the second step of the process. In this second step, a query search aims to rank the query search results such that a user may find the result they are searching for within a reduced set of query results. It is this second step where PageRank finds it's applicability.

Personalized PageRank can be viewed as a specific sub-case of the general PageRank where the algorithm attempts to exploit additional knowledge contained in the conditional information of a specified user. This is best illustrated in an example; consider a social network such as Facebook. If a specific user is searching for someone named "John", the ranking algorithm would be inefficient if it simply returned a list of Johns that exist in the network. If instead, the algorithm put into consideration information conditional on the specific user such as the user's location or users named John who are connected to friends of the user, the algorithm can greatly improve the accuracy of it's rankings. From a mathematical perspective, Personalized PageRank begins by consideration of nodes within a notion of distance away from the User's personal node. From this perspective, a user conducting a query search on the network is herself represented as a node in the network from which the search is centralized around.

In the years since the inception of Personalized PageRank, the theoretical details surrounding these ranking algorithms has exploded. Of the plethora of ideas surrounding Personalized PageRank, the specific research thread concerning its ranking distribution has found a great deal of interest [12][13][14]. Early implementations which are still commonly seen today assume Ergodic Markov Chains [13] satisfying convenient classical probability theorems such as the classical central limit theorem. As an alternative yet equalivalent standpoint, one can view Personalized PageRank as the endpoint distribution of a collection of random walks each of which has a strictly positive exponentially-distributed number of steps. However, many of the assumptions surrounding these results may be inaccurate. Probability theory surrounding classical random walks on mathematical structures such as lattices and manifolds is well understood while analysis on complex networks is still a growing subject [15]. Of the many possibilities around this problem, **a centerpiece of the problem may be focused on the relationship between the random walk's step length and the network's structure, including the impact induced by changing the step length distribution.** The goal of the research behind this paper consists of the beginning stages of testing the impact on changing the walk's step distribution on the final ranking distribution provided by Personalized PageRank. In particular, we are interested in differences compared to classical Personalized PageRank induced by changing the step length distribution to a power-law, as well as convergence results under these changes.

The results of our research are presented as follows. In section 2, we begin with a short expository of related work in this field, followed by an introduction to Zipf's law and the interesting implications surrounding this distribution. Furthermore, we give motivation for implenting this distribution within Personalized PageRank. Next, section 3 introduces the three datasets to be considered in our analysis.

To begin our analysis, we test the convergence assumptions of classical PageRank on our modified PageRank algorithm in section 4. In an effort to be both broad and concrete, we analyze our convergence results using two different metrics: Absolute Score Differences and Mean Absolute Deviation. We conduct this analysis on our Facebook dataset, being the largest dataset in our analysis and thus giving the most consistent answers. On top of this, we include convergence results over a broad range of parameter values with direct comparisons to the classical algorithm.

Section 5 includes an in-depth analysis of the differences between the rankings provided by our modified PageRank algorithm in comparison to the classical version. In analyzing the differences, we develop two different metrics to measure the differences in the ranking distributions. Results of our analysis are mixed; although we find qualitative evidence for dependency on the walk's step distribution, the evidence is not strong enough to make decisive conclusions. Finally, in a concluding section, we briefly review our results and provide motivation for further study into this subject.

II. Related Work

The original PageRank algorithm including Personalized PageRank was introduced in [1]. A more recent analysis utilizing the assumptions made on the random walk's step distribution is conducted in [16]. The research conducted in [16] makes explicit the connection between the algebraic equation for PageRank and it's interpretation as a random walk with geometrically distributed step length. A major piece of research that was the basis behind our ideas comes from [17]. The authors in [17] conduct extensive simulations on the empirical distribution constructed by applying PageRank on artificially-created network structures. In their analysis, they vary the in-degree of the graph and analyze it's impact on the empiracle distribution of PageRank's ranking distribution. They find that by creating a graph with in-degree following a power-law with exponent in the range of 2.1-2.2 leads to non-exponential decrease of PageRank's ranking distribution for specific instances. In particular, if the parameter of PageRank (α) is around .85, they find the ranking distribution to follow power-law behavior. They fit a Pareto distribution - a common power-law distribution - on the empirical distribution with very minimal error. Due to α typically being chosen between $.85$ and $.90$ in practice, this illustrates the potential for power-laws to exist within PageRank ranking distributions.

Based on the results in [17] mentioned above, this leads one to ask the following question: does changing PageRank's random walk distribution to a power-law lead to similar results? One can view this question as the converse of the questions posed in [17]; as opposed to altering the structure of the network while maintaining the usual geometrically-distributed random walk step-length, we chose to instead alter the random walk step-length to a power-law. Unlike the analysis conducted in [17] however, we conducted our analysis on real datasets representing networks that are pervasive in our everyday lives.

The Pareto distribution, also known as "Zipf's law" or the Zipfian distribution has it's origins in the slightly related field of word occurances [18]. In short, a set of random variables are said to take values according to Zipf's law if the probability of occurance is inversely proportional to the magnitude of the value. Given the prevalence of this distribution in frequency of words, one may be led to believe there could also be a similar relationship in query searching on networks. The possiblity of this relationship is precisely what led us to using the Zipfian distribution in this research.

III. Datasets Under Consideration

The first data set under consideration consists of friend's lists from Facebook borrowed from the SNAP database [19]. This dataset, being the largest one under consideration in our research, is represented as a network consisting of 4039 nodes with 88234 edges. Due to the immense computational complexity in analyzing these datasets, we restrict our convergence analysis in section 4 to this dataset.

The second dataset under consideration represents innovation between physicians located in towns in Illinois, Peoria, Bloomington, Quincy and Galesburg. The dataset maintained and freely accessible from the Konect database [20]. In this dataset, each node represents a physician while an edge represents communication between physicians, deemed an "innovation". This is a relatively small dataset in comparison to the facebook dataset: it consists of a total of 241 nodes (physicians) with 923 edges (innovations).

The final dataset we analyze in this paper represents a network of interactions between members of Zarchary's Karate Club, also freely available from the Konect Database[21]. In this case, each node represents one of the members of the club, while edges represent interactions between members in the club. This network is by far the smallest we consider in our research, made up of a total of 34 nodes with 78 edges.

IV. Random Walk Convergence

The classical Personalized PageRank algorithm consists of solving a linear matrix-algebraic problem:

$$\pi_s = \alpha\pi + (1-\alpha)\mathbf{1} \tag{1}$$

Where π_s represents the stationary distribution generated by the PageRank algorithm. Equivalently, this problem can be viewed as a system of random walks each taking a Geometrically distributed number of steps with parameter α: $\sim Geo(\alpha)$. We adopt this viewpoint of a system of random walks throughout the rest of this paper.

The difference between PageRank and Personalized Pagerank derives completely from a difference in the initial condition: for PageRank, the typical assumption is that each random walk begins with equal probability at any of the nodes within the network. On the other hand, Personalized PageRank assumes each walk begins at a pre-specified node with probability 1.

A. Modification of PageRank

To modify our Personalized PageRank algorithm, we choose to draw a Zipfian-distributed $\sim Zipf(a)$ number of steps for each walk, each beginning at the pre-specified beginning node. To begin our convergence analysis, we first analyze the Sum of Absolute PageRank Score Differences for increasing number of walks:

$$\sum_{k=1}^{m} |Score_k(n) - Score_k(n-1)| \quad (2)$$

Where our function $Score_k(n)$ represents the empirical probability of node k out of the m nodes in the network. In our case, $Score_k(n)$ is simply computed by taking the number of walks that ended at node k and dividing by the total number of walks:

$$\frac{\sum_n \chi_k}{n} \quad (3)$$

Where χ represents the indicator function.

B. Convergence Under Sum of Absolute PageRank Score Differences

Figures 1 and 2 illustrate the results of our experiment for differing parameter values of the Geometric(α) (respectively Zipfian(a)) distribution.

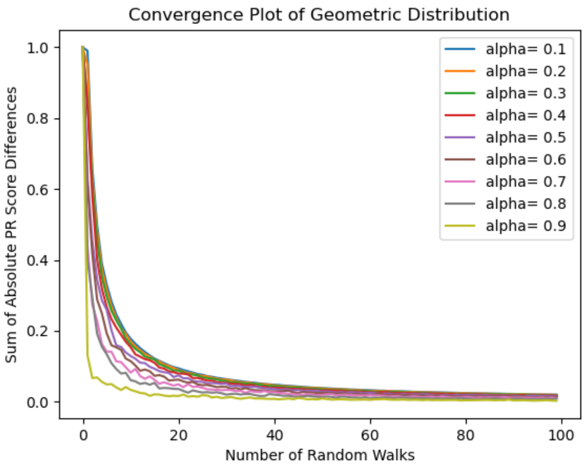


Fig. 1: Convergence of Personalized PageRank using Geometric step distribution under Sum of Absolute PageRank Score Differences by differing α as number of walks n increases

Results of these experiments indicate the ranking distribution for both the classical Personalized PageRank Algorithm as well as our modified version are converging to stationary distributions as the number of random walks increase. However, it is clear from Figures 1 and 2 that for general parameter values, the Geometrically-distributed step length walks converge to a stationary distribution slower than our Zipfian-distributed step length walks. To illustrate this point, figures 3 and 4 provide information about the number of walks required such that the Sum of Absolute PageRank Score Differences falls beneath a threshold of $.004$. Overall, our takeaway from these figures is that in both cases, the algorithm is converging to a stationary rank-score distribution. Futhermore, and somewhat surprisingly, it appears that by using a Zipfian distribution we find considerable speed-up in convergence to a stationary distribution.

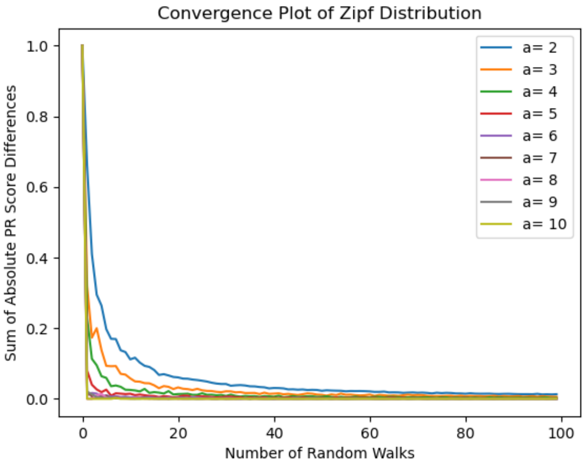


Fig. 2: Convergence of Personalized PageRank using Zipfian step distribution under Sum of Absolute PageRank Score Differences by differing a as number of walks n increases

C. Convergence Under Mean Absolute Deviation (MAD)

Despite the positive convergence results illustrated by the sum of absolute PageRank score differences, we wish to further analyze convergence by analyzing the variability of our estimates as the number of walks increase. In order to find our MAD estimates, we repeat our experiment T times, where for each node we find the average score $\bar{\pi}_k$ across all of the experiments:

$$\bar{\pi}_k = \sum_{i=1}^{T} \frac{\frac{\sum_n \chi_k}{n}}{T} \quad (4)$$

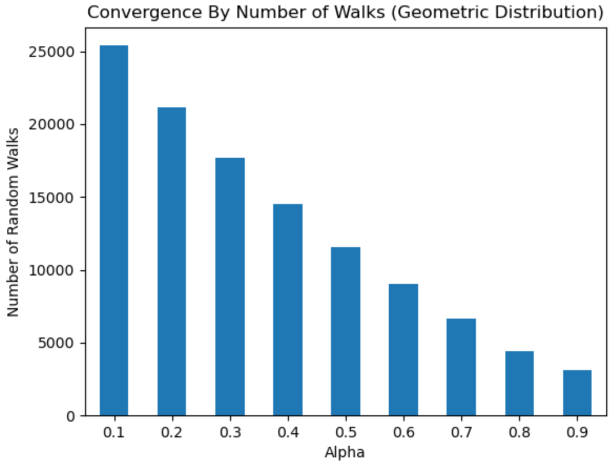


Fig. 3: Number of walks required for convergence under the threshold value .004 using the Geometrically-distributed step distribution

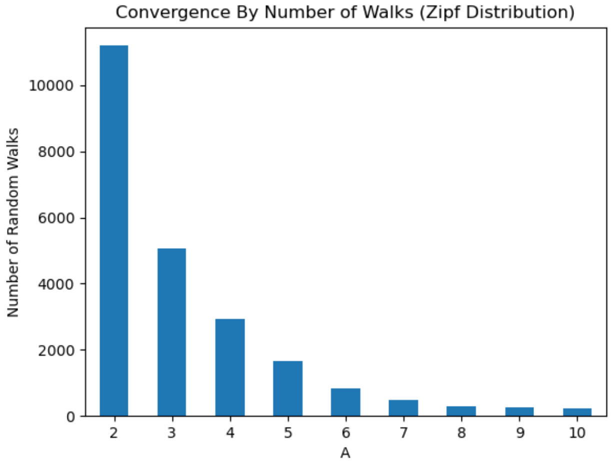


Fig. 4: Number of walks required for convergence under the threshold value .004+ using the Zipfian-distributed step distribution

Under the typical assumptions of PageRank, we are assuming that as the number of walks increases, the variability is decreasing to zero:

$$lim_{n\to\infty}\frac{1}{n}|\bar{\pi_k}(n)-\bar{\pi_k}(n-1)|=0 \tag{5}$$

Figures 5 and 6 illustrate MAD for differing values of the parameters of the Geometric and Zipfian distributed step distributions. Both plots indicate that the MAD is decreasing towards zero, hence providing evidence that using either distribution is leading to a stationary distribution.

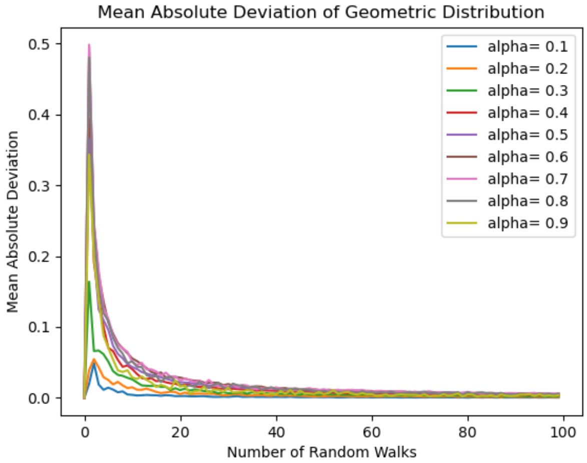


Fig. 5: Convergence of Personalized PageRank using Geometrically-distributed step length under Mean Absolute Deviation by differing α as number of walks n increases

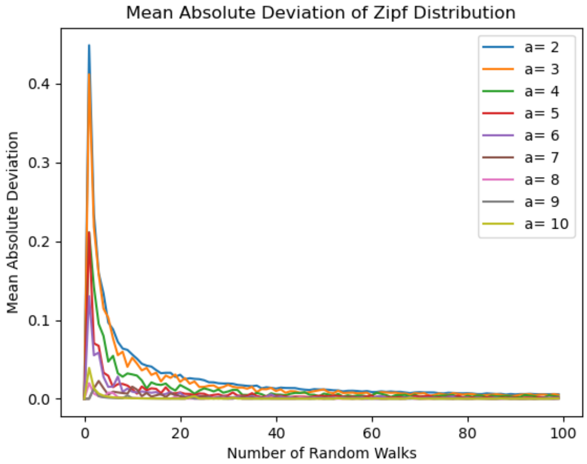


Fig. 6: Convergence of Personalized PageRank using Zipfian-distributed step length under Mean Absolute Deviation by differing a as number of walks n increases

V. Differences Between Classical and Modified PageRank Distributions

In section 4, we found evidence that both the classical Personalized PageRank, as well as our Modified PageRank algorithm is converging to a stationary distribution on our Facebook dataset. In this section we illustrate the differences in ranks provided by PageRank under the Geometrically-distributed step distribution in comparison to the Zipfian-distributed step distribution across the three different networks introduced in section 3. Before conducting the analysis, we expected to find the ranks provided by the Zipfian-distributed step distribution to differ dramatically from the classical algorithm; one would expect a significant score going further out into the network from the initial node when using a power-law distribution.

A. *Measures of Distributional Difference*

In an attempt to analyze the differences across the distributions provided by the classical algorithm against our modified algorithm, we designed two very different measures of difference. The first, which we denote as the **rank-based measure**, involves three steps. The first is finding the top-5 ranked nodes found using the Geometric-distributed step random walk. Secondly, we find the ranks of the nodes found in the first step under the Zipfian-distributed step random walk. Finally, in the third step, we take the absolute differences between each of these five ranks, and sum the values. More rigorously, denote $Score_{\alpha}(k)$ as the empirical probability of node k under Geometric-distributed step random walks with parameter α. As is standard in PageRank, we rank the nodes as a partial ordering over these score values: define $G_1, ..., G_m$ as the first, second, ..., m-th ranked nodes:

$$Score_{\alpha}(G_1) \geq Score_{\alpha}(G_2), ..., \geq Score_{\alpha}(G_m). \tag{6}$$

Next, we define the function $rank_{\alpha}(k)$ as the value corresponding to the rank designated by the above construction. Analogously, define $Score_a(k)$ as the empirical probability of node k under Zipfian-distributed step random walks with parameter a, and $rank_a(k)$ as the corresponding rank values. With the above definitions in place, we define our rank-based measure as follows:

$$D_{rank}(\alpha, a) = \sum_{k=1}^{5} |k - rank_a(G_k)|. \tag{7}$$

The second, and more direct measure of difference involves summing the absolute differences in empirical distribution scores directly. We denote this measure as our **Score-based measure**:

$$D_{score}(\alpha, a) = \sum_{k=1}^{m} |Score_{\alpha}(k) - Score_a(k)|. \tag{8}$$

B. *Distributional Differences On The Facebook Dataset*

We begin our analysis of the differences between the Zipfian-distributed step distribution and the Geometric-distributed step distribution first on the Facebook network dataset. Figures 7 and 8 illustrate the rank-scores provided to each node in the form of a heat-map. Due to the immense size of this network, it can be quite difficult to gain a qualitative understanding of these plots. However, it is clear that our modified Personalized PageRank algorithm tends to spread it's scoring out over a larger number of nodes for general parameter values in comparison to the classical algorithm.

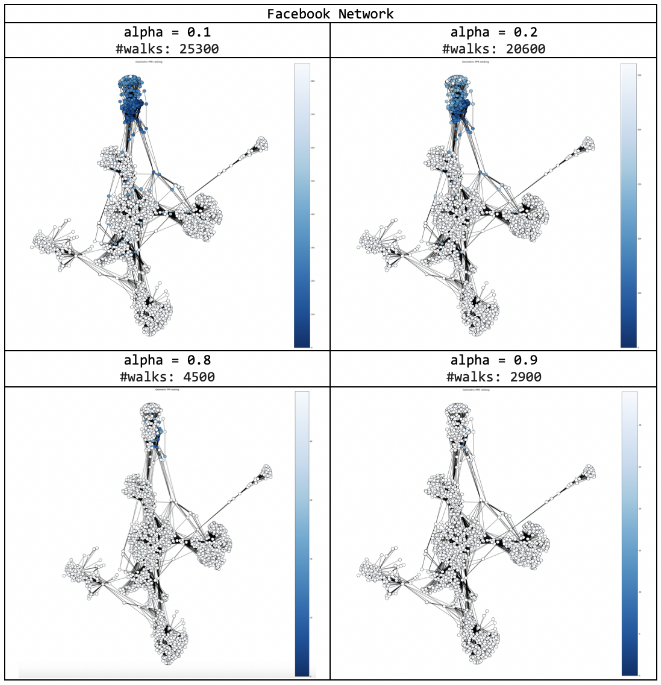


Fig. 7

Next, we analyze the distributional differences by our two measures of distributional difference. Figures 9 and 10 provide the measure values for several parameter values corresponding to our two random walk step distributions. From figure 9, we can see that as the parameter a of the Zipfian distribution increases, $D_{rank}(\alpha, a)$ decreases across all values of the parameter α of the geometric distribution. This appears to tell us what our intuition

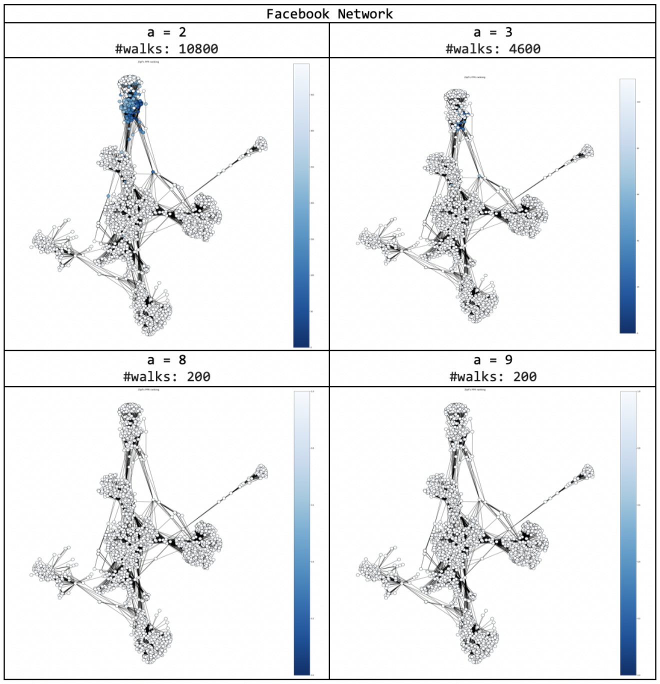


Fig. 8

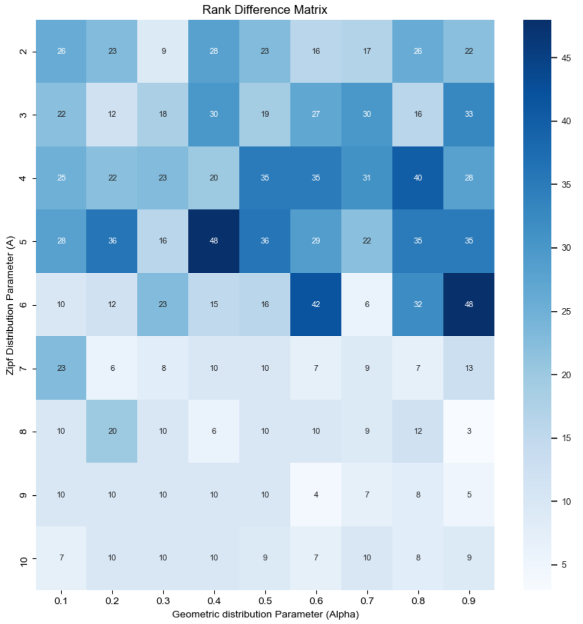


Fig. 9

would have assumed: as a increases, the effect of power-law tails is dampened to the point that the Zipfian-based walks give nearly the same rankings as Geometric-distributed walks.

Despite our results under the rank-based measure, our score-based measure $D_{score}(\alpha, a)$ appears to tell a very different story. Figure 10 appears to illustrate that as we increase α while simultaneously decrease a, the two algorithm's estimates are converging to the same values. **We consider this to be a significant result, and refer the reader to our concluding section for an in-depth discussion**.

C. Distributional Differences On The Physician Dataset

As opposed to our Facebook dataset, our Physician Dataset corresponds to a much smaller network, allowing for ease of interpretation of plots corresponding to distributional heat maps. Figure 11 shows the distributional heat maps for four different values of α corresponding to the classical algorithm, while figure 12 illustrates the heat maps under our modified algorithm.

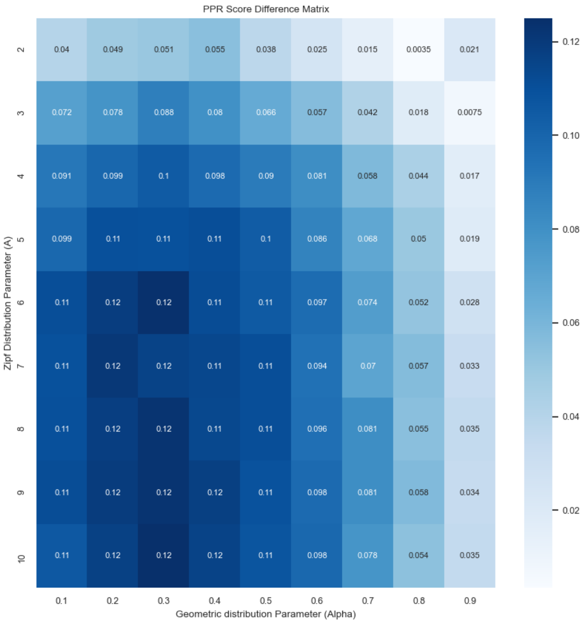


Fig. 10

Similarly to the Facebook dataset, these plots appear to be similar in the two algorithms for low values of both α and a, while illustrating the extreme concentration of the distributions for high values. Next, we analyze the dataset with our two

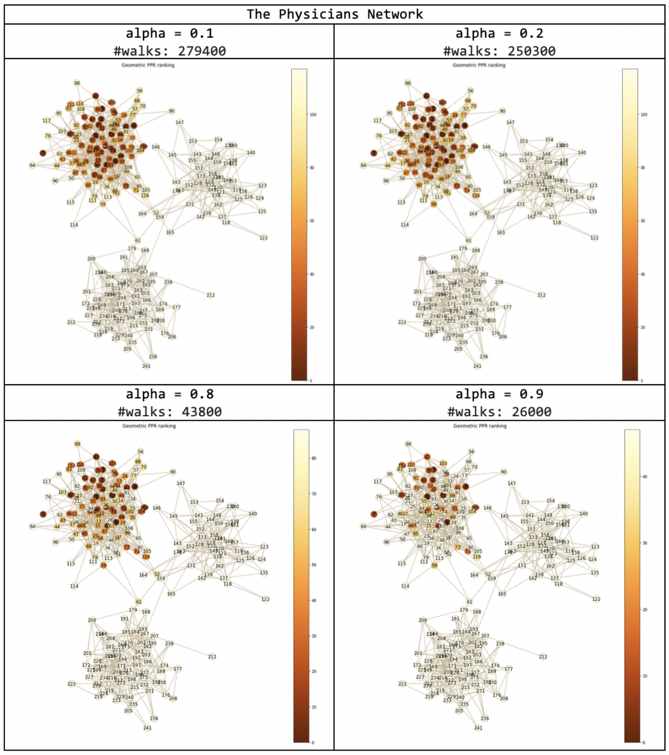


Fig. 11

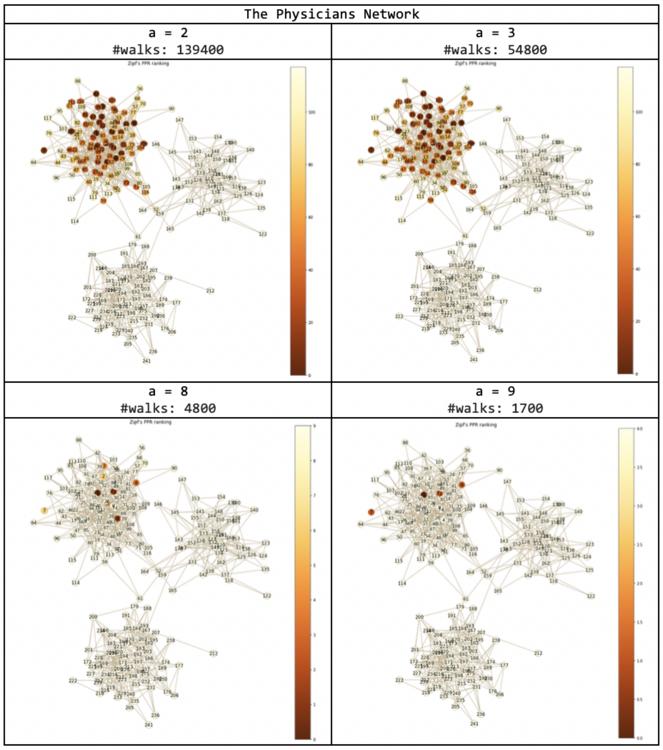


Fig. 12

measures of distributional differences in figures 13 and 14.

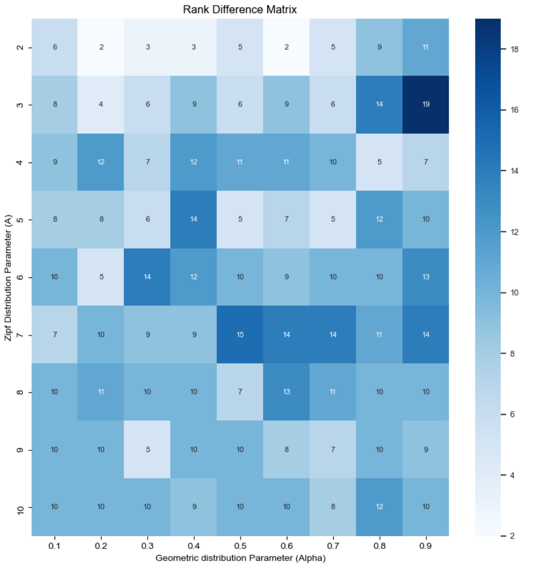


Fig. 13

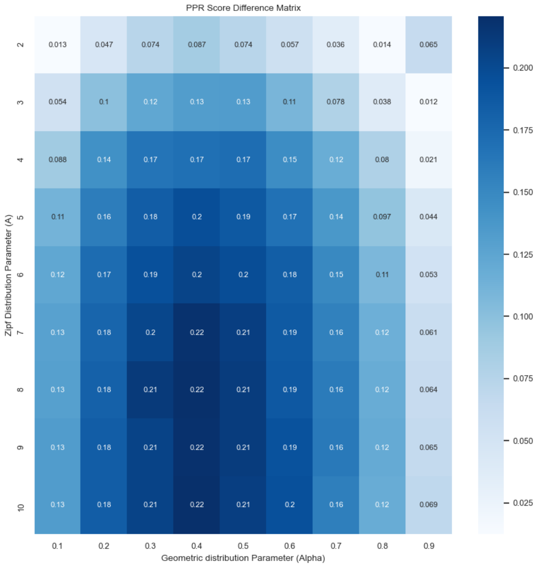


Fig. 14

A significant departure from our results found in the Facebook dataset is found with our rank-based measure: in this case, there appears to be no direct relationship between increasing or decreasing either of the parameters. However, our score-based measure remains consistent with our findings on the Facebook dataset: as we increase α while simultaneously decrease a, the two algorithm's estimates are converging to the same values.

D. Distributional Differences On The Karate Dataset

Our final dataset is by far our smallest dataset, consisting of only 34 nodes and 78 edges. The small size of this dataset allows for ease of interpretation of distributional heat maps provided in figures 15 and 16. Qualitatively, our interpretation is analogous to our findings in both the Facebook and Physician datasets.

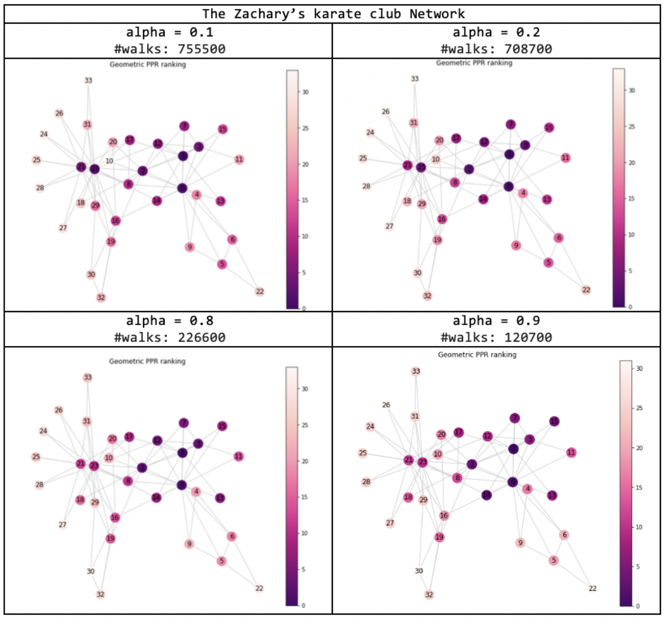


Fig. 15

Finally, we analyze the Karate Dataset under our two measures of distributional difference. Figures 17 and 18 provide tables for our computed measure values of rank-based and score-based differences, respectively.

Similarly to the Physician dataset, there appears to be no direct relationship between increasing or decreasing the parameter values of the step distributions in terms of our rank-based measure. If we could conclude anything from this figure, it would be that the opposite of our findings in the Facebook dataset: increasing the Zipfian distribution parameter a appears to decrease the difference between these two distributions for all values of α. However, unlike our rank-based measure, the score-based measure displayed in Figure 18 appear to show the same relationship found in the previous two datasets: as we increase α while simultaneously decrease a, the two algorithm's estimates are converging to the same values.

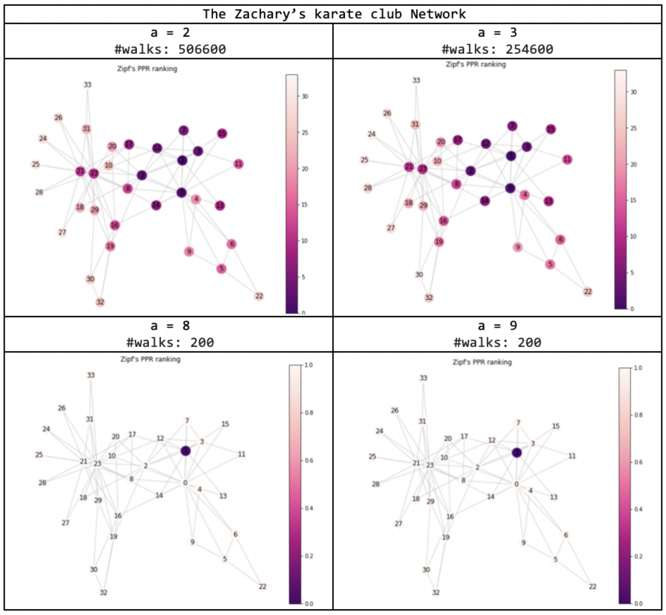
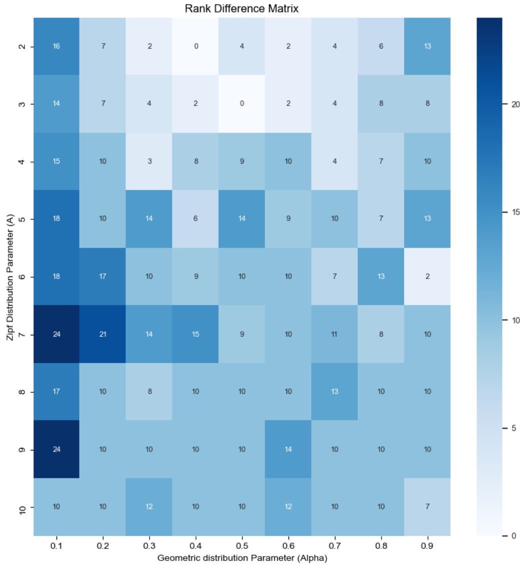


Fig. 16



	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	16	7	2	0	4	2	4	6	13
3	14	7	4	2	0	2	4	8	8
4	15	10	3	8	9	10	4	7	10
5	18	10	14	6	14	9	10	7	13
6	18	17	10	9	10	10	7	13	2
7	24	21	14	15	9	10	11	8	10
8	17	10	8	10	10	10	13	10	10
9	24	10	10	10	10	14	10	10	10
10	10	10	12	10	10	12	10	10	7

Fig. 17

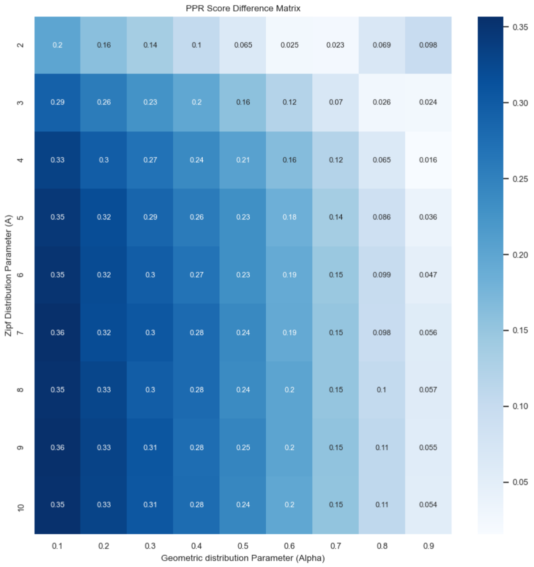


Fig. 18

VI. Conclusion and Future Directions

In setting out on conducting the research portrayed in this paper, we had strong expectations of what we might find. In particular, we had expected to find strong differences between our Modified PageRank algorithm and the classical PageRank algorithm in their rankings - while still maintaining convergence to a stationary distribution in both cases. In terms of their ranking differences, these expectations stem from knowing how different a power-law distribution can be from a classical exponential-law distribution. As for convergence, due to the finite size of the networks under consideration, we had predicted we would find convergence regardless of the step-size distribution.

While the analysis conducted in section 4 seems to provide strong evidence in favor of our expectation, the analysis of distributional differences in section 5 was both a surprise and an exciting development. To illustrate the importance of this result, we recall some of the results found in [17]. Here, as noted in section 2, the authors constructed artificial networks with power-law in-degree distributions over a wide range of power-law exponents. In their findings, they highlight the unique behavior found when the in-degree power-law exponent was in the range of $2.1 - 2.2$, while maintaining the PageRank parameter α within $.85 - .9$. For these specific parameter values, they find the distribution of PageRank scores distinctively follow power-law behavior fitted by the Zipfian distribution. **In short, they found that the distribution of the classical PageRank scores is strongly influenced by the structure of the network**. Our research objective was to approach this concept from the opposite direction: for a given network, can changing the step-size implicit in the PageRank algorithm induce a similar result? From the results illustrated in sections 4 and 5, we have provided strong evidence of deviation in distribution, while still maintaining the property of stationarity of the distribution of scores.

The summary of our analysis outlined above provides an immense range of possible future projects. From a theoretical perspective, the most obvious question we have in mind revolves around the interdependence apparent between the network structure and the random walk's step distribution. The effect of this interdependence appears to have a strong influence the resulting scores provided by the PageRank algorithm. If information on this interdependence can be gained, one may be able to improve upon the PageRank algorithm to provide optimal rankings. The size of these datasets are growing exponentially as time goes on, driving the need for ideas to continue to improve upon the PageRank Algorithm.

VII. Appendix

Code is available here:

```
https://drive.google.com/open?
id=10-_tOTxaE1hxVzN5j7vazL7x5YCZ0J99
```

REFERENCES

[1] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web.", Stanford InfoLab, Tech. Rep., 1999.

[2] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine", 1998.

[3] M. Richardson, A. Prakash, and E. Brill, "Beyond pagerank: Machine learning for static ranking", in *Proceedings of the 15th International Conference on World Wide Web*, ser. WWW '06, Edinburgh, Scotland: Association for Computing Machinery, 2006, pp. 707–715, ISBN: 1595933239. DOI: 10.1145/1135777.1135881.

[4] F. Zhu, Y. Fang, K. C.-C. Chang, and J. Ying, "Incremental and accuracy-aware personalized pagerank through scheduled approximation", 2013.

[5] L. Jiang, B. Ge, W. Xiao, and M. Gao, "Bbs opinion leader mining based on an improved pagerank algorithm using mapreduce", in *2013 Chinese Automation Congress*, 2013, pp. 392–396.

[6] S. Kamvar, T. Haveliwala, and G. Golub, "Adaptive methods for the computation of pagerank", *Linear Algebra and its Applications*, vol. 386, pp. 51–65, 2004.

[7] S. Kamvar, T. Haveliwala, C. Manning, and G. Golub, "Exploiting the block structure of the web for computing pagerank", Stanford InfoLab, Technical Report 2003-17, 2003. [Online]. Available: http://ilpubs.stanford.edu:8090/579/.

[8] C. Platzer and S. Dustdar, "A vector space search engine for web services", in *Third European Conference on Web Services (ECOWS'05)*, 2005, 9 pp.-.

[9] N. *, *How many websites are there around the world? [2020]*, Feb. 2020. [Online]. Available: https://www.millforbusiness.com/how-many-websites-are-there/.

[10] J. Clement, *Facebook: Number of monthly active users worldwide 2008-2019*, 2019.

[11] *53 incredible facebook statistics and facts*. [Online]. Available: https://www.brandwatch.com/blog/facebook-statistics/.

[12] H. Kao and S. Lin, "A fast pagerank convergence method based on the cluster prediction", in *IEEE/WIC/ACM International Conference on Web Intelligence (WI'07)*, 2007, pp. 593–599.

[13] L. Breyer, "Markovian page ranking distributions: Some theory and simulations", *Preprint*, 2002.

[14] S. Bai, L. Chen, and J. Wu, "Heterogeneous information network based ranking and clustering of mobile apps",

[15] N. Masuda, M. A. Porter, and R. Lambiotte, "Random walks and diffusion on networks", *Physics Reports*, vol. 716-717, pp. 1–58, Nov. 2017, ISSN: 0370-1573. DOI: 10.1016/j.physrep.2017.07.007.

[16] P. Lofgren, S. Banerjee, and A. Goel, "Bidirectional pagerank estimation: From average-case to worst-case", in *International Workshop on Algorithms and Models for the Web-Graph*, Springer, 2015, pp. 164–176.

[17] L. Becchetti and C. Castillo, "The distribution of pagerank follows a power-law only for particular values of the damping factor", in *Proceedings of the 15th international conference on World Wide Web*, 2006, pp. 941–942.

[18] A. Gelbukh and G. Sidorov, "Zipf and heaps laws' coefficients depend on language", in *International Conference on Intelligent Text Processing and Computational Linguistics*, Springer, 2001, pp. 332–335.

[19] *Social circles: Facebook*. [Online]. Available: https://snap.stanford.edu/data/ego-Facebook.html.

[20] *Physicians*. [Online]. Available: http://konect.uni-koblenz.de/networks/moreno_innovation.

[21] *Zachary karate club*. [Online]. Available: http://konect.uni-koblenz.de/networks/ucidata-zachary.