

数理実習 B(離散手法) レポート

学生証番号：03-190615

氏名：工藤龍

メールアドレス：ron233f@gmail.com

2019 年 6 月 20 日

1 概要

研究室のホームページの役割には様々なものがある。その中でも重要なものが、これから研究室を決める学生や、共同研究先を探している企業の人などが見る場合である。そのような人たちが求めているものは、研究室がどのような研究をしていて、どんな人が所属しているかなどの基本的な情報を素早くわかりやすい形で入手することである。

これを鑑みると、この目的における「より良い」ウェブサイトとは、余分な情報はあまり含まず、重要な情報がアクセスしやすいサイトであるということができる。

今回の解析では、計数工学科システム情報工学コースの 8 個の研究室を比較し、どの研究室のウェブサイトがより理想的な形になっているかを考察した。

2 解析手法

各研究室の全てのページに関してそれぞれのサイト内でのページランクを算出した。その後、それをヒストグラムとして可視化し、それぞれのサイトの特徴を考察した。ページランク算出の際のダンピング係数は 0.85 とした。

3 結果

ページランクを各研究室のウェブサイトの全てのページについて算出した結果、次の図 1 から図 8 までのヒストグラムのようになった。

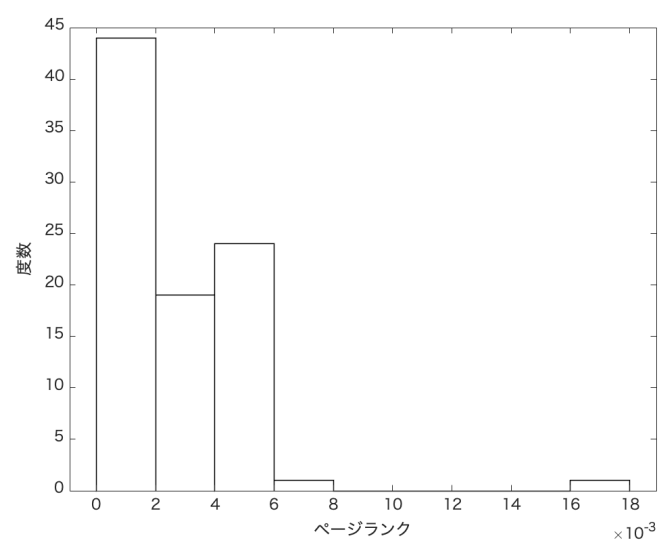


図 1: 第 1 研究室のページランクの分布

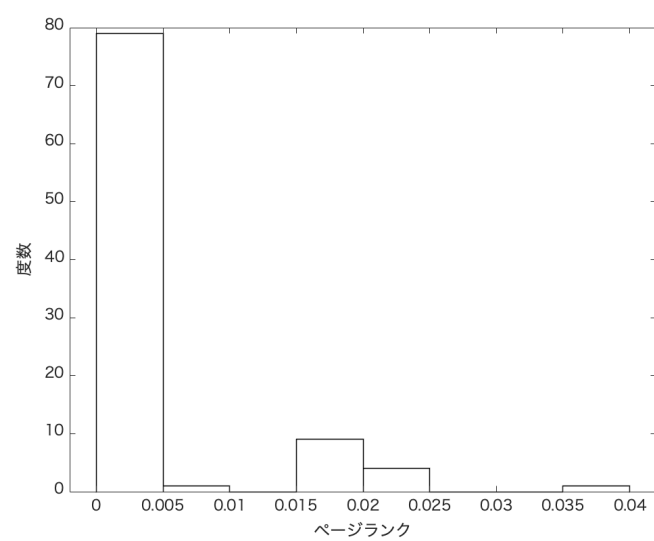


図 2: 第 2 研究室のページランクの分布

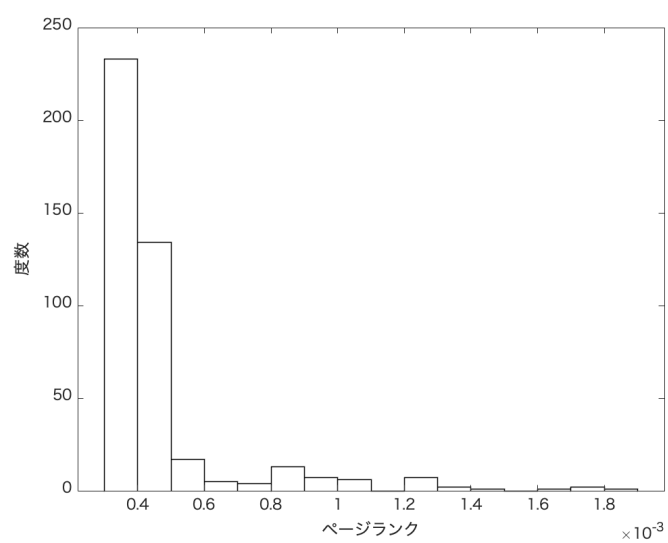


図 3: 第 3 研究室のページランクの分布

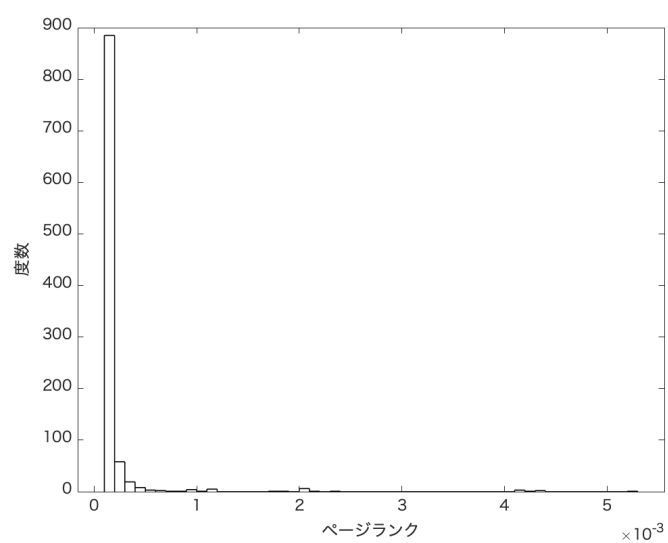


図 4: 第 4 研究室のページランクの分布

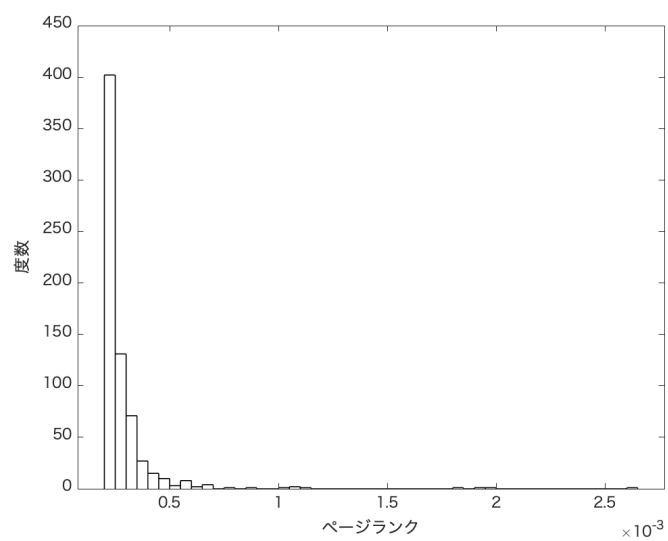


図 5: 第 5 研究室のページランクの分布

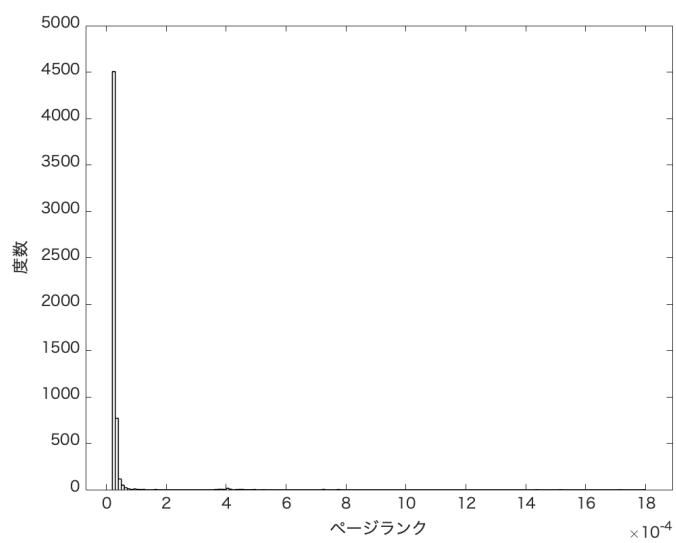


図 6: 第 6 研究室のページランクの分布

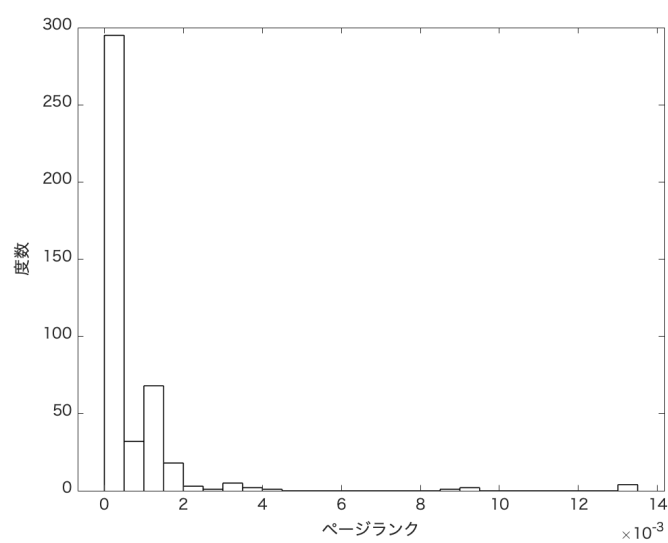


図 7: 第 7 研究室のページランクの分布

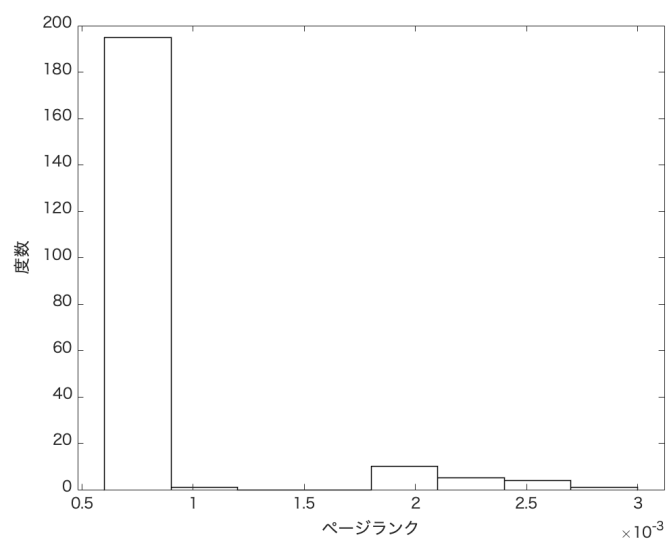


図 8: 先端研 田研究室的ページランクの分布

4 考察

ヒストグラムを全体としてみると、まずわかることとして研究室によってヒストグラムの形状が全く異なることがあげられる。ページランクの低いものがどのサイトも多くなっている点では共通しているが、その割合は大きく異なっている。研究室ごとにコンテンツの加え方・記載している情報と捨象している情報などの運用方針が異なることが推察される。

第2研究室を見ると、ページランク最大のもののページランクが唯一0.35を超えていることがわかる。これは、ある一つのページに多くのリンクが繋がっていて、かつそのページから多くのページにリンクが繋がっていることを意味する。相対的にページランクが小さいものの割合が小さいため、重要な情報がまとまっていると予想される。ゆえに学生や企業の人にとっては「良い」サイトと言えるだろう。

また、もう一つ特徴的な研究室が第6研究室である。ここは、ページランクの小さいものの数が異常に多く、相対的にページランクが大きいものの数が少ない。これは、たくさんのページに大量の情報を保持し、ある種のアーカイブ的なものとしてウェブサイトを運用していることが見て取れる。一方、学生や企業など外部の人から見ると、情報過多な状態となりアクセスしにくい可能性もある。

今回の解析では、研究室のウェブサイトの目的を、外部の人が研究室の概要を理解するためと絞って考えた。しかし、実際には研究室のウェブサイトは多種多様な役割を果たしている。例えば、研究室の内部の人のためのアーカイブという役割などが考えられる。また、学生と企業の人では求める情報も異なるはずである。学生はおそらく研究室の研究内容だけではなく雰囲気などの要素も重視するであろう。一方の企業では、研究内容はもちろんのこと過去に他の企業と連携した経験があるか、といった要素も見erkかもしれない。

このように、どんな目的を設定するかによってウェブサイトの適切な解析方法は異なってくるものであり、今回の解析はその一例にすぎない。他の目的ではどんな解析が適切かを考察するのは今後の課題としたい。