

2019-07-12

# Measuring Similarity Between Mathematics Research and Its References Using Bibliographic Coupling

Hansen, Samuel

<https://hdl.handle.net/2027.42/150102>

<http://creativecommons.org/licenses/by-nc/4.0/>

---

*Downloaded from Deep Blue, University of Michigan's institutional repository*

# MEASURING SIMILARITY BETWEEN MATHEMATICS RESEARCH AND ITS REFERENCES USING BIBLIOGRAPHIC COUPLING

## WHO?

Samuel Hansen  
Mathematics & Statistics Librarian  
University of Michigan  
hansensm@umich.edu

## WHAT?

An analysis of the development of mathematics research overtime, with a focus on similarity between publication-reference and citation-publication pairs.

## WHY?

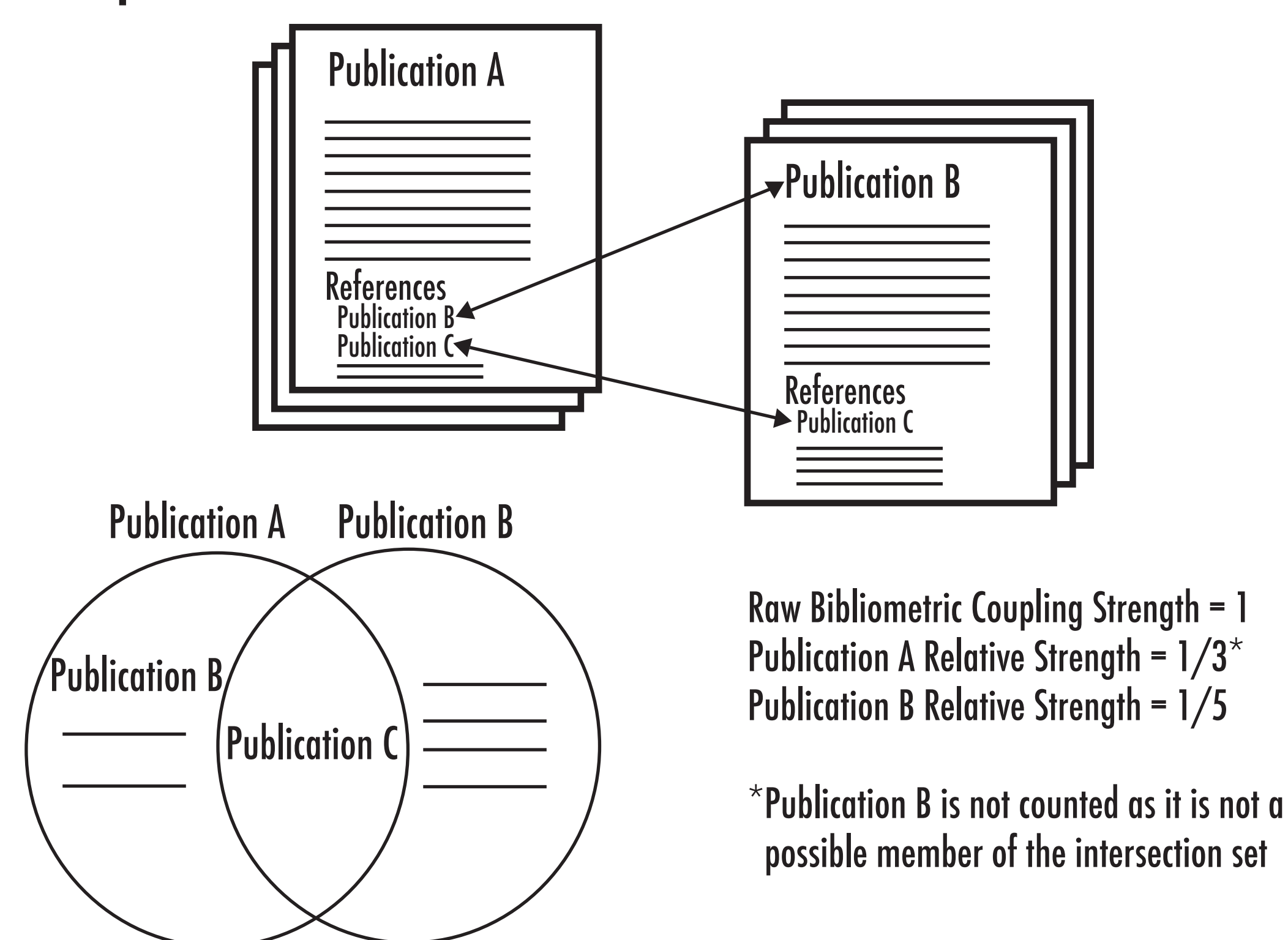
In order to understand mathematics it is important to develop knowledge of how it has grown and the ways it builds upon itself.

## USING?

The data used is from the Clarivate Web of Science citation database, 1900-2017. All records with Web of Science category designations Mathematics; Mathematics, Applied Mathematics; and Mathematics, Interdisciplinary Applications were included in the analysis.

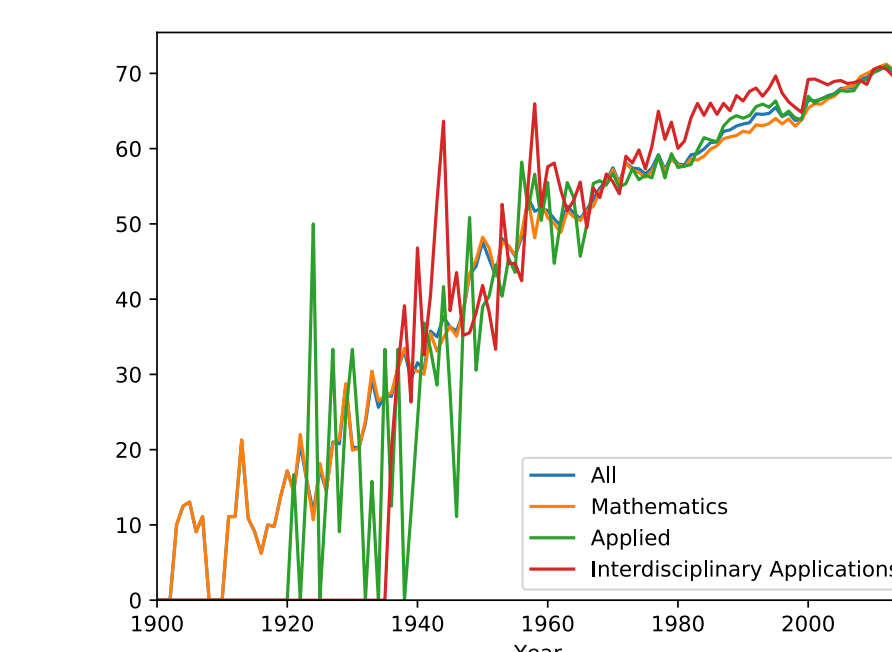
## HOW?

Similarity between publication pairs is measured using raw and proportional Bibliometric Coupling strength. Raw strength is the number of publications in the intersection of reference sets. The proportional strengths takes the raw counts and divide them by sizes of each publication's total reference set.

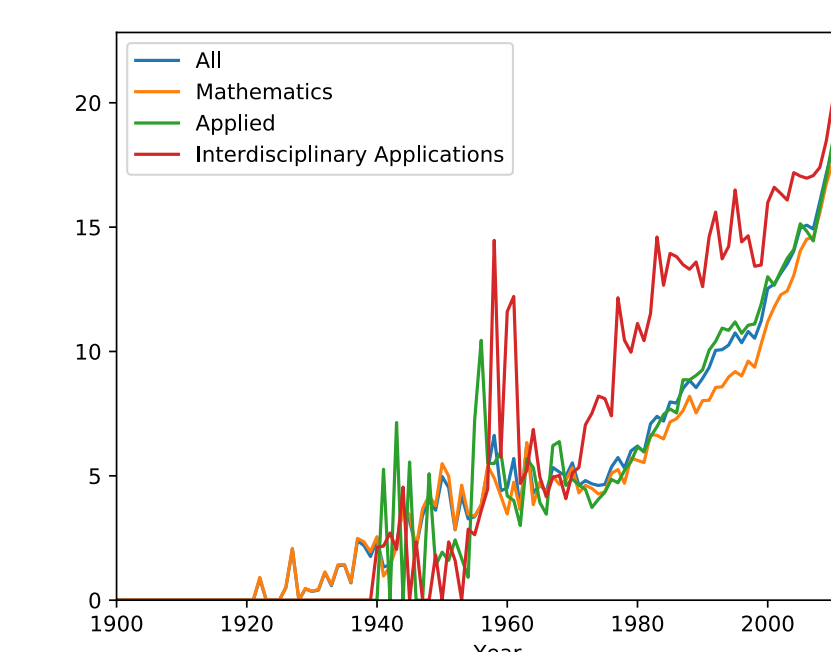


## PUBLICATION-REFERENCE PAIRS

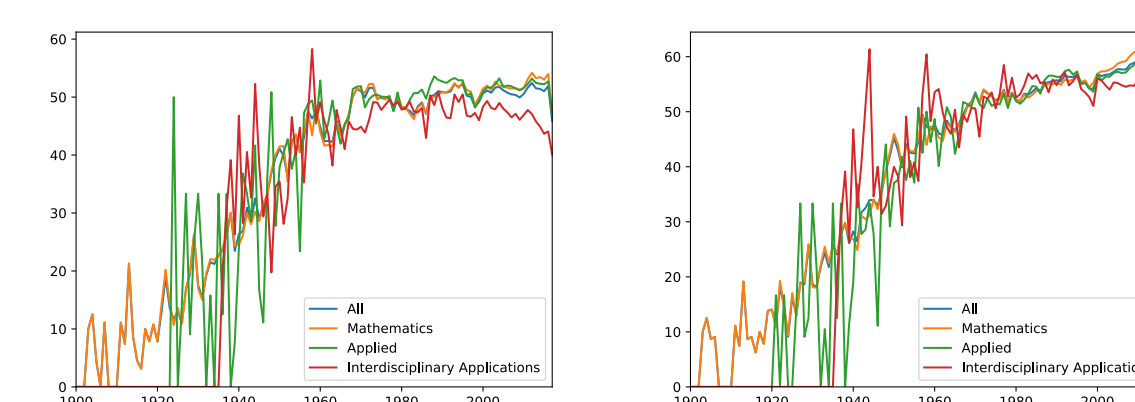
Category	Publication-Reference Pairs	At least 1 Shared Reference (% of total)	At least 3 (% of total)	At least 5 (% of total)	10 or more (% of total)
Mathematics	4802134	3221604 (67.09%)	1443842 (30.07%)	681683 (14.20%)	133506 (2.38%)
Applied	3765012	2579284 (68.51%)	1228600 (32.63%)	617010 (16.38%)	136962 (3.64%)
Interdisciplinary	918937	631977 (68.77%)	317735 (34.58%)	168611 (18.33%)	44386 (4.83%)
<b>Total</b>	<b>7901392</b>	<b>5348251 (67.69%)</b>	<b>2481515 (31.41%)</b>	<b>12216547 (15.40%)</b>	<b>263331 (3.33%)</b>



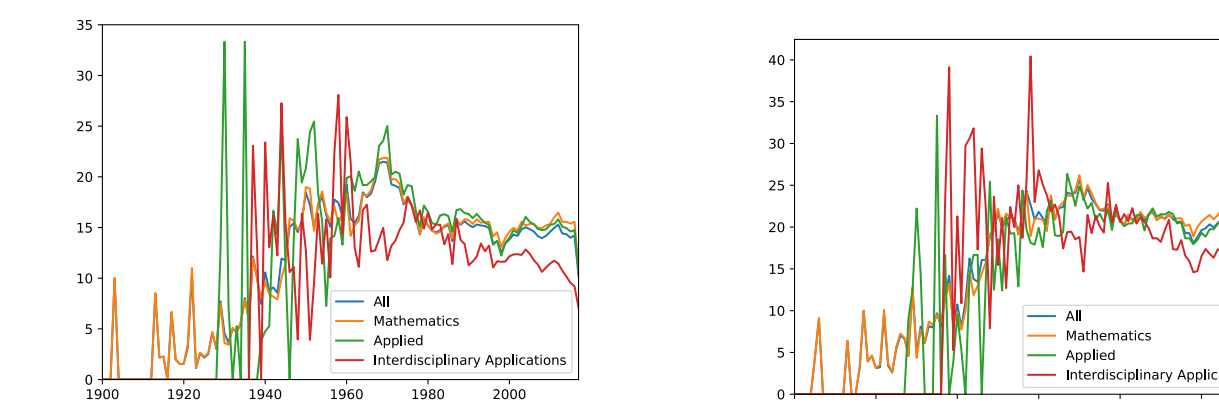
Percentage of Publication-Reference pairs sharing at least 1 reference over time



Percentage of Publication-Reference pairs sharing at least 5 references over time



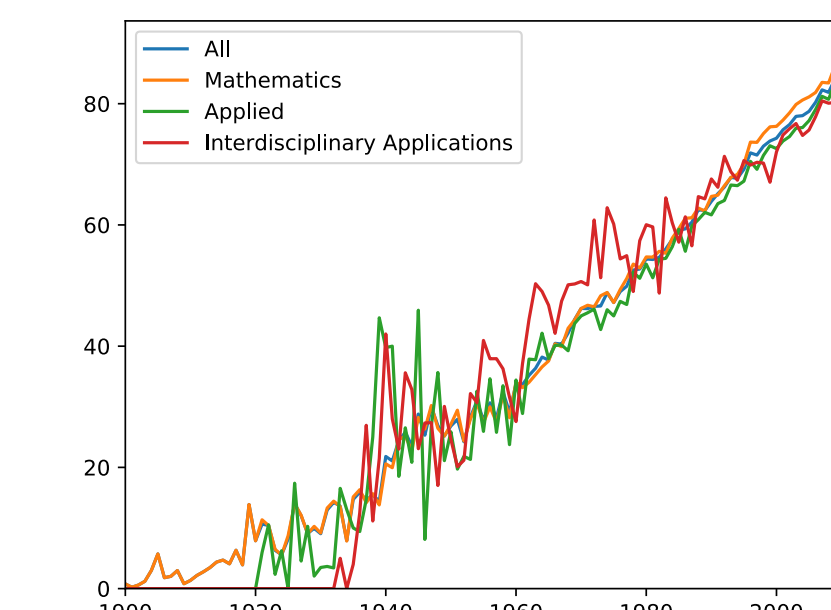
Publication Reference  
Percentage of Publication-Reference Pairs with relative coupling strength greater than 5% over time



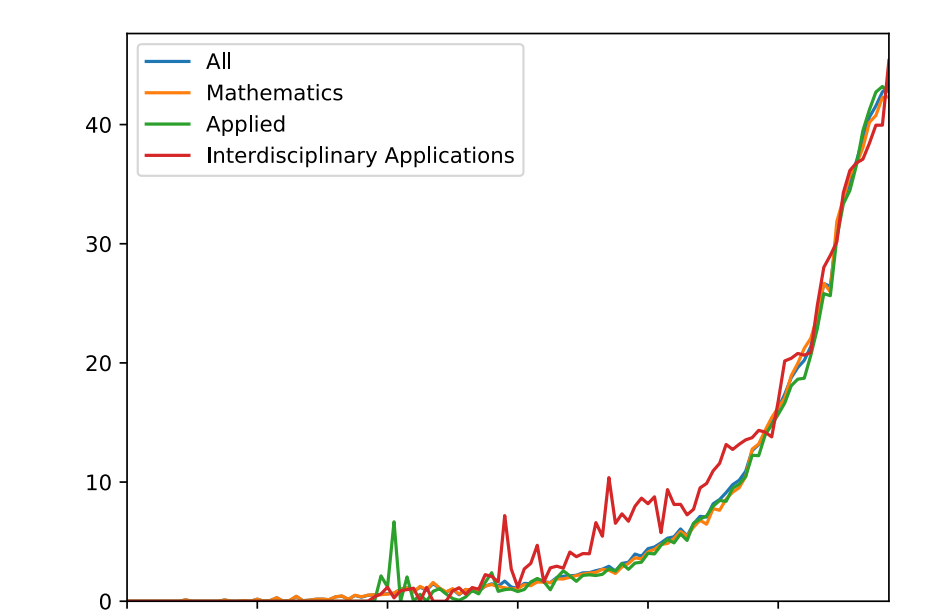
Publication Reference  
Percentage of Publication-Reference Pairs with relative coupling strength greater than 20% over time

## CITATION-PUBLICATION PAIRS

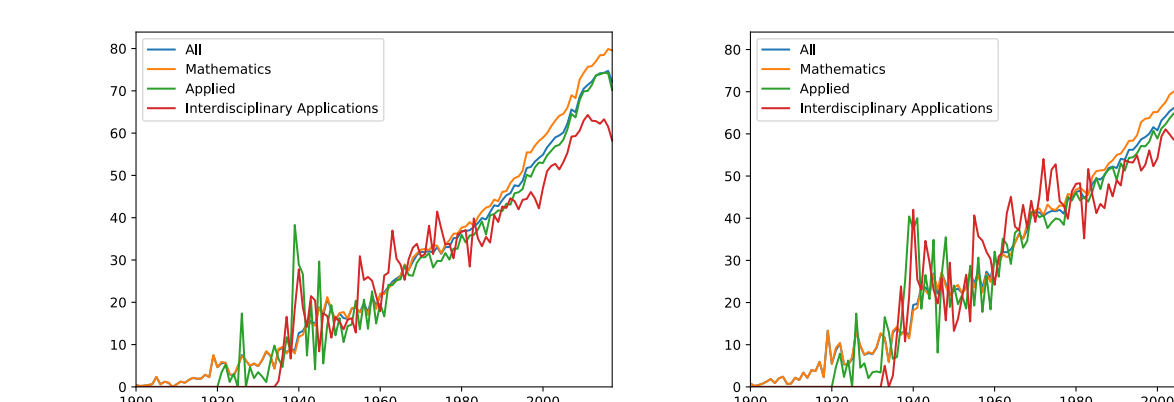
Category	Citation-Publication Pairs	At least 1 Shared Reference (% of total)	At least 3 (% of total)	At least 5 (% of total)	10 or more (% of total)
Mathematics	4967056	3273539 (65.91%)	1442886 (29.05%)	675081 (13.59%)	131860 (2.65%)
Applied	3390481	2391880 (70.61%)	1175232 (35.66%)	599808 (17.69%)	136182 (4.02%)
Interdisciplinary	815449	585437 (71.79%)	306705 (37.61%)	165914 (20.35%)	44696 (5.48%)
<b>Total</b>	<b>7919132</b>	<b>5361943 (67.71%)</b>	<b>2489505 (31.41%)</b>	<b>1221186 (15.42%)</b>	<b>264558 (3.34%)</b>



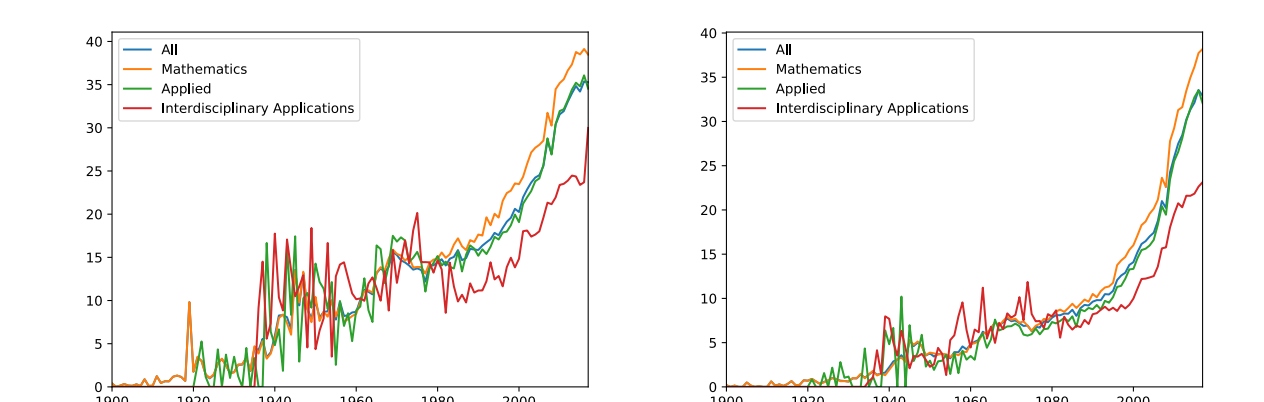
Percentage of Citation-Publication pairs sharing at least 1 reference over time



Percentage of Citation-Publication pairs sharing at least 5 references over time



Publication Citation  
Percentage of Citation-Publication Pairs with relative coupling strength greater than 5% over time



Publication Citation  
Percentage of Citation-Publication Pairs with relative coupling strength greater than 20% over time

## HIGHLY CITED PUBLICATIONS

Category	Citation-Publication Pairs	At least 1 Shared Reference (% of total)	At least 3 (% of total)	At least 5 (% of total)	10 or more (% of total)
Mathematics	936977	429699 (45.86%)	123869 (13.22%)	45808 (4.89%)	6730 (0.72%)
Applied	673396	373582 (55.48%)	131497 (19.53%)	53714 (7.98%)	8788 (1.31%)
Interdisciplinary	225874	142328 (63.01%)	60078 (26.60%)	26721 (11.83%)	154 (0.80%)
<b>Total</b>	<b>1634021</b>	<b>838287 (51.30%)</b>	<b>280299 (17.15%)</b>	<b>112472 (6.88%)</b>	<b>18367 (1.12%)</b>

## OBSERVATIONS

- While raw coupling strength for Publication-Reference pairs has increased over time, especially for higher strengths, relative coupling strength on the other hand shows a flattening for lower relative strengths and a decrease for higher one since the 1960s. This is likely related to the dramatic increase in references per publication, from a median around 7 to over 20 since 1960.
- The relative coupling strength of Citation-Publication pairs shows a steady increase over time, but evidence indicates this may be caused by citing publications sharing fewer references with older publications, behavior which would drive down relative strength as a publication ages.
- There is a noticeable decrease in the raw coupling strength of Citation-Publication pairs when publications with more than 100 citations are considered. There are many possible causes of such behavior, including popular papers serving as stand-ins for the publications they reference or highly cited publications being cited due to popularity instead of applicability.
- The more applied publications had higher raw coupling strength, while the more pure mathematics publications had higher relative strengths. Since applied publications by their nature will reference a broader range of publications it is logical for pure mathematics publications to be more similar relative to their reference lists.

For references and a work-in-progress paper please visit [tinyurl.com/shbibpapersla2019](http://tinyurl.com/shbibpapersla2019)