



Pinpointing Influence in *Pinterest*

Panagiotis Liakos¹ Katia Papakonstantinou¹ Michael Sioutis²
Konstantinos Tsakalozos³ Alex Delis¹

¹University of Athens – ²Université d'Artois, CRIL – ³Canonical Group Ltd.

2nd International Workshop on Social Influence Analysis
(co-located with IJCAI 2016)
New York City, July 9th, 2016

Motivation

- The extreme growth of online social networks enables us to study influence patterns at scale.
- We want to answer if there exist certain individuals with the power to **affect** their social contacts and **convince** them to **buy a product** or **adopt a political idea**.

Identifying influential individuals

allows for cost-effective viral marketing techniques to **increase brand awareness** or even **sway the public opinion!**

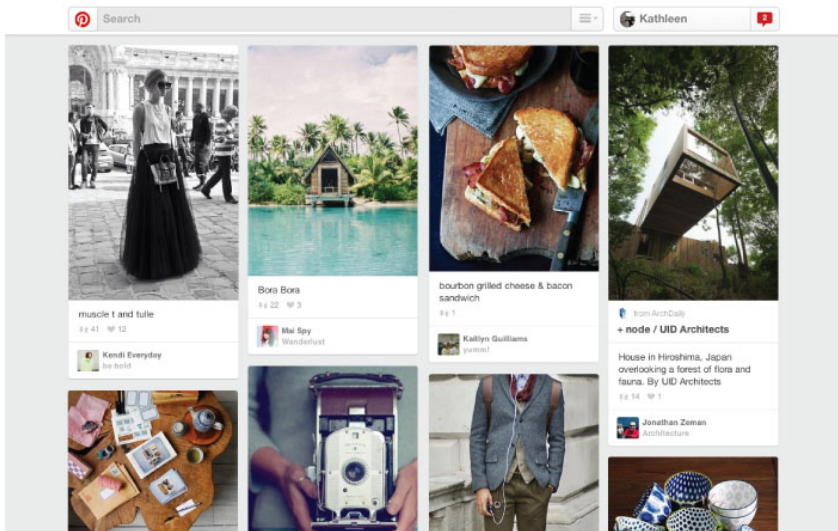
Studies on **Twitter** [CHBG10] reveal that topological measures such as indegree **fail** to capture the influential strength of users.

We perform an in-depth empirical analysis on *Pinterest* and seek to answer:

- Is the finding of [CHBG10] true across **other online social networks** as well, and **to what extent**?
- Does the use of **PageRank** [LSMW98] allow for a **better estimation** of a user's influential power?

What is Pinterest?

Pinterest is a visual bookmarking tool that helps you discover and save creative ideas.



Pinterest lets you:

Pin something to a board
and come back to it later to learn more.

The screenshot shows a Pinterest board interface. At the top, there is a search bar with the text "Search" and a magnifying glass icon. To the right of the search bar is the Pinterest logo. Further right is a user profile section for "Morgan" with a plus sign and a camera icon. Below this is the board title "Grub" in a large, bold font, followed by the subtitle "All the things I want to cook, bake and eat." Below the subtitle is a row of controls: a profile picture and name "Morgan Keys", an "Edit Board" button, a "Send Board" button with a share icon, and two statistics: "32 Pins" and "276 Followers". The main content area is a grid of pins. On the left is a dashed box with a plus sign and the text "Add a Pin". The first pin shows soft pretzels with the text "from Spoonful: 15 Must-Have Soft Pretzel Recipes" and "These soft pretzels look delicious and perfect for our potluck." The second pin shows a Snickers Semifreddo Cake with the text "Snickers Semifreddo Cake | Bakers Royale" and "Pinned from bakersroyale.com". The third pin shows Caramelised Onion & Goat Cheese Tartlets with the text "Caramelised Onion & Goat Cheese Tartlets with Balsamic Syrup" and "Pinned from whatkatieate.com".

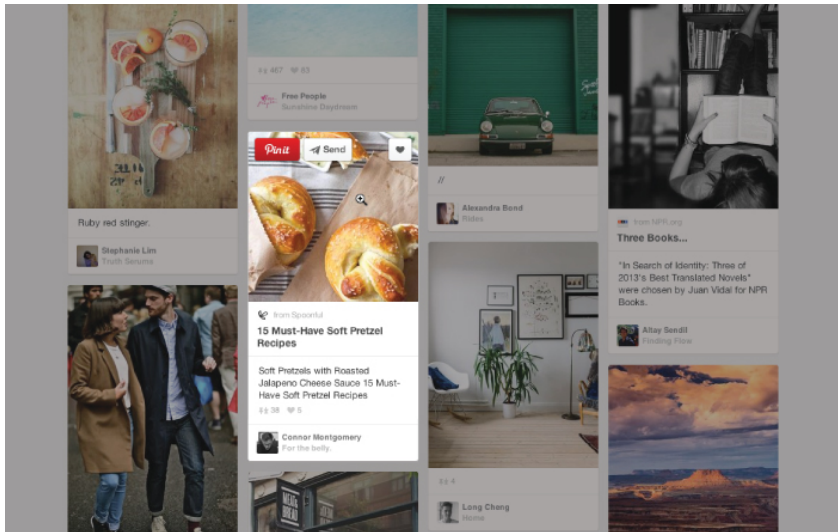
Pinterest lets you:

Follow people whose taste you admire to receive their pins in your home feed.

The screenshot shows the Pinterest interface for a user named Becky Stoneman. At the top, there is a search bar with the Pinterest logo and the word "Search", a menu icon, and a profile header for "Kathleen" with a notification badge showing "2". Below this is the profile header for "Becky Stoneman", which includes her profile picture, a "Follow" button, and a settings gear icon. A statistics bar below the header shows: 32 Boards, 1,960 Pins, 0 Likes, 88 Followers, and 128 Following. The main content area features three boards: "Styyyy" with a blue and white striped top (264 Pins), "Dress It Up" with a colorful geometric dress (118 Pins), and "Live to Romp!" with a white lace skirt (33 Pins). Each board has a "Follow" button at the bottom. The boards are displayed in a grid with a main image and a row of four smaller thumbnail images below it.

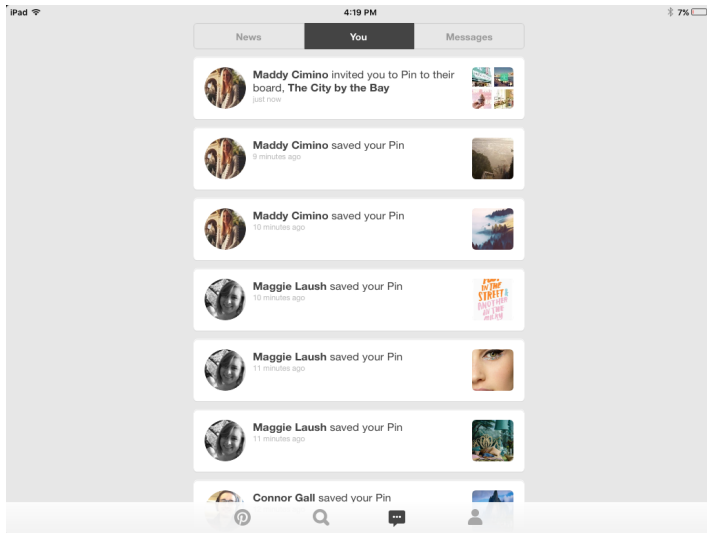
Pinterest lets you:

Repin or Like pins of others.



Pinterest lets you:

See how others **interact** with your **pins**.



Why Pinterest?

Pinterest stands out for many reasons:

- It was the fastest site to surpass 10,000,000 monthly active users.
- It has more than 100,000,000 monthly active users.
- Its vast majority of users are female.

Pinterest has attracted significant commercial attention:

Users tend to create **digital shopping lists** of products they are interested in **buying**.

Therefore, businesses **invest** in creating compelling boards in *Pinterest* to **increase their revenue**.

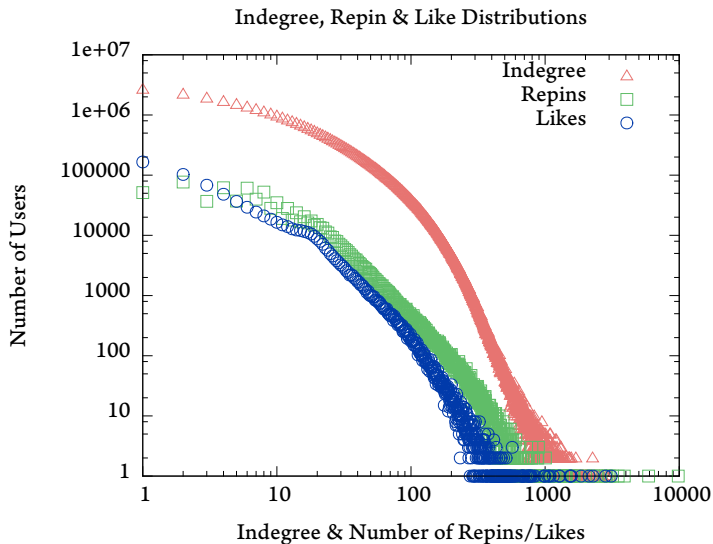
Definition of Influence in Pinterest

- **Indegree influence**: the number of followers of a user directly indicates the *size of the audience* of that user.
- **PageRank influence**: the PageRank of a user indicates the *strength of her influence on her followers*.
- **Like influence**: the number of likes containing one's name indicates the ability of that user to generate *popular* content.
- **Repin influence**: the number of repins containing one's name indicates the ability of that user to generate content with *pass-along value*.

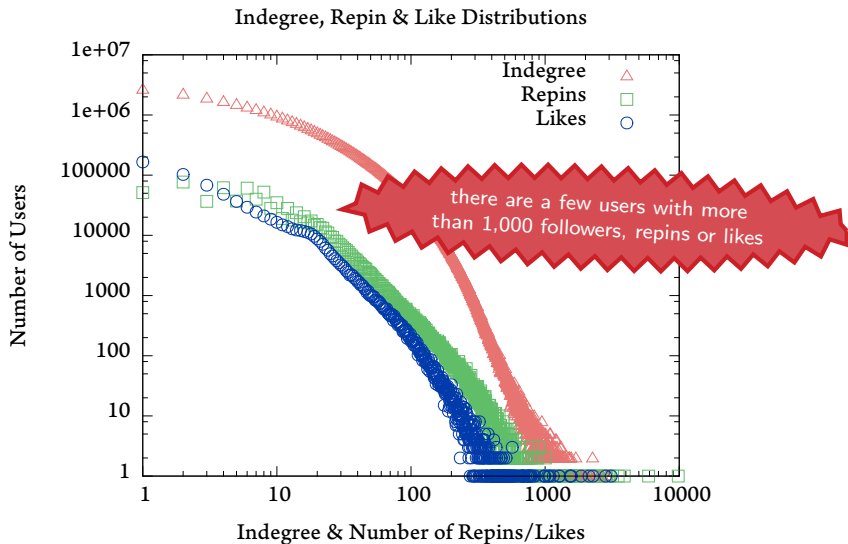
- **Dataset [ZSS⁺14, ZSSS13]:**
 - 36,198,633 users
 - 983,520,986 social ties
 - 18,957,340 repins
 - 9,066,973 likes
- **PageRank Execution:**
 - Dell PowerEdge R630 server with an Intel[®] Xeon[®] E5-2630 v3, 2.40 GHz processor, and 256 GB of RAM
 - Deployed an Apache Hadoop 2.7.1 cluster using Juju¹
 - Run PageRank as an Apache Giraph process

¹<https://jujucharms.com/big-data>

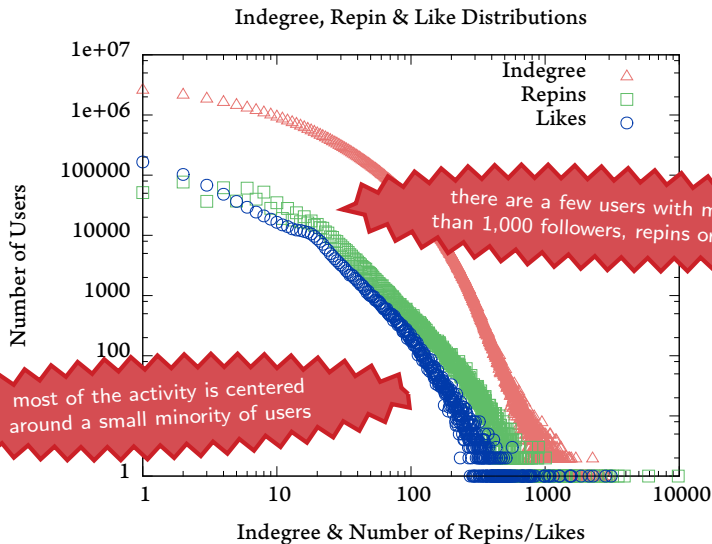
Distribution of indegree and received repins/likes



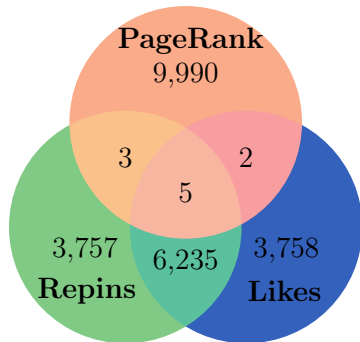
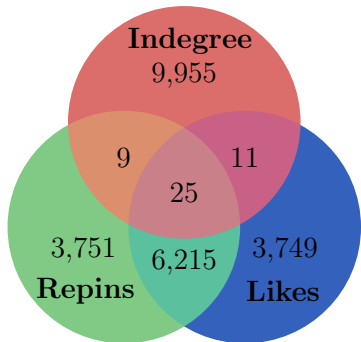
Distribution of indegree and received repins/likes



Distribution of indegree and received repins/likes



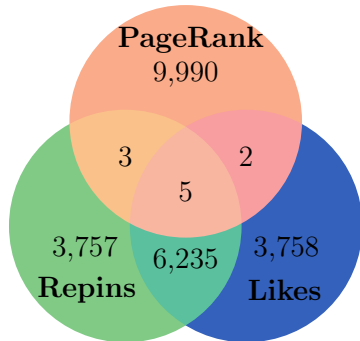
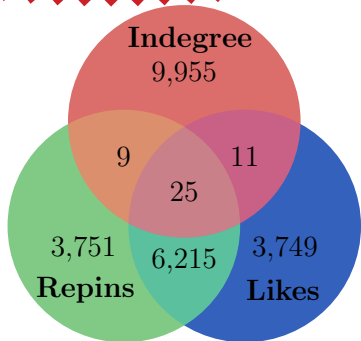
Overlap of Top-Ranked Users



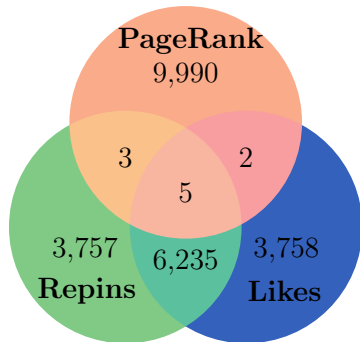
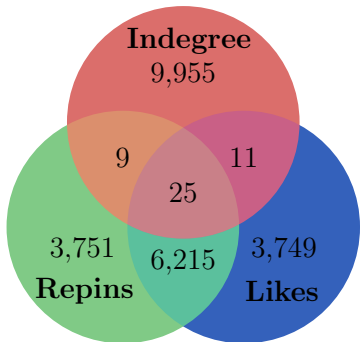
Overlap of top-10,000 users ranked by the measures of influence under consideration

Overlap of Top-Ranked Users

the overlap of indegree with both repins and likes is marginal



Overlap of Top-Ranked Users



the overlap of PageRank with repins and likes is also insignificant

Overlap of Top-Ranked Users

Indegree
9,955

PageRank
9,990

Hints of very weak correlation
of the *indegree* or *PageRank* of users
with the frequency they receive *repins* and *likes*.

Repins

Likes

Repins

Likes

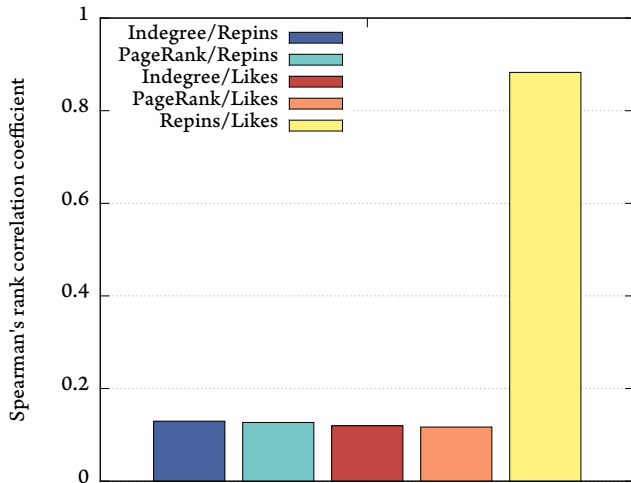
Comparing Influence Measures

- For all measures of influence:
 - We assigned the rank of 1 to the most influential user, and increased the rank as we proceeded to less influential users.
 - Identical values were each assigned fractional ranks equal to the average of the positions in the ascending order of the values.
- We used Spearman's rank correlation coefficient ρ to examine whether two rankings covary.

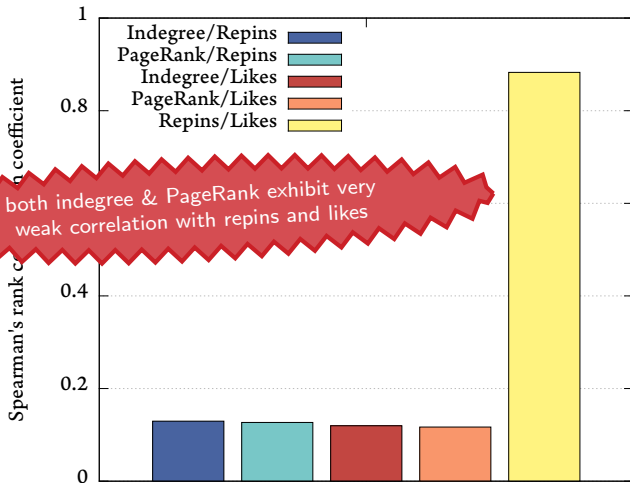
$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

where $d_i = rg(X_i) - rg(Y_i)$ is the difference between the two ranks of user i , and n is the total number of users.

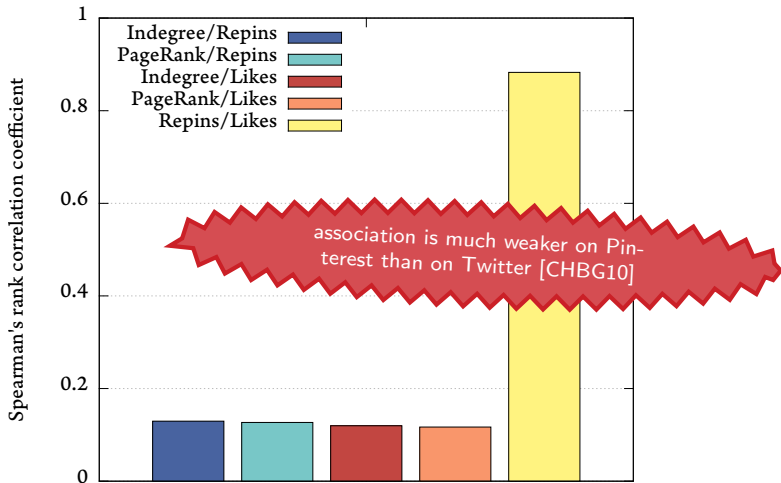
Rank correlation for all users



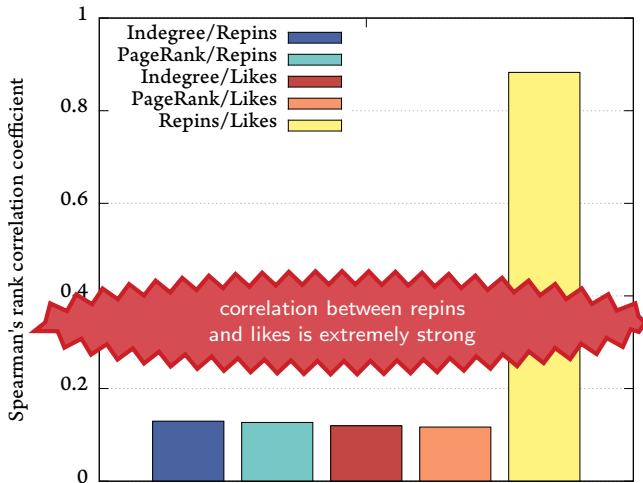
Rank correlation for all users



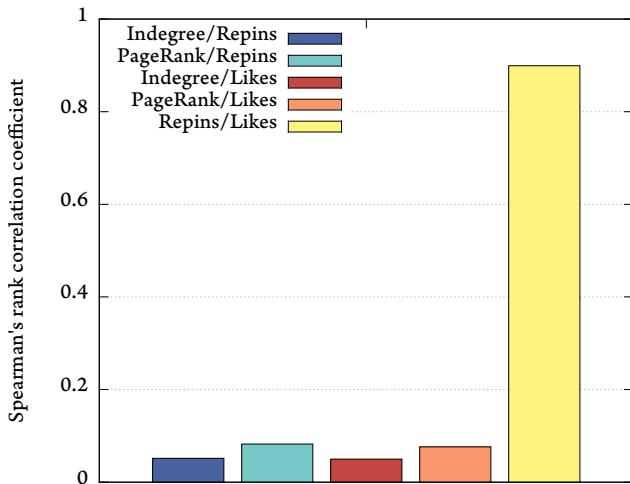
Rank correlation for all users



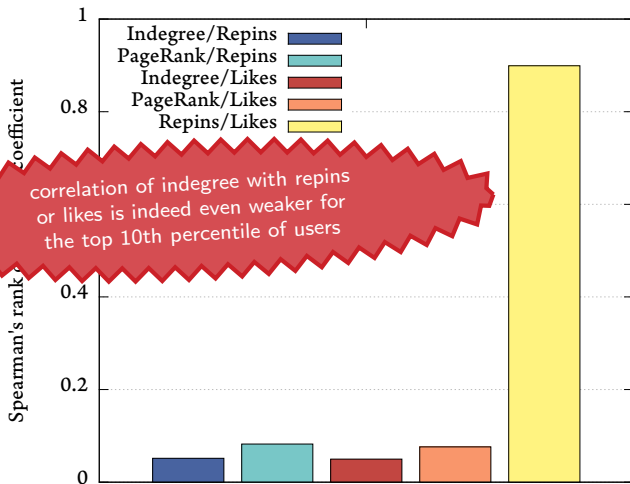
Rank correlation for all users



Rank correlation for the top 10th percentile of users

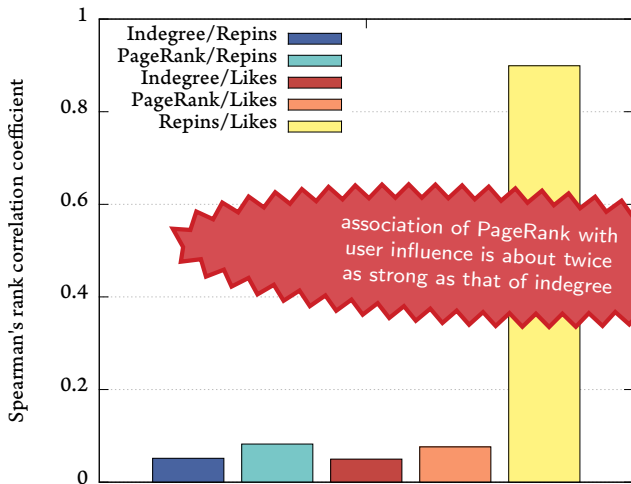


Rank correlation for the top 10th percentile of users

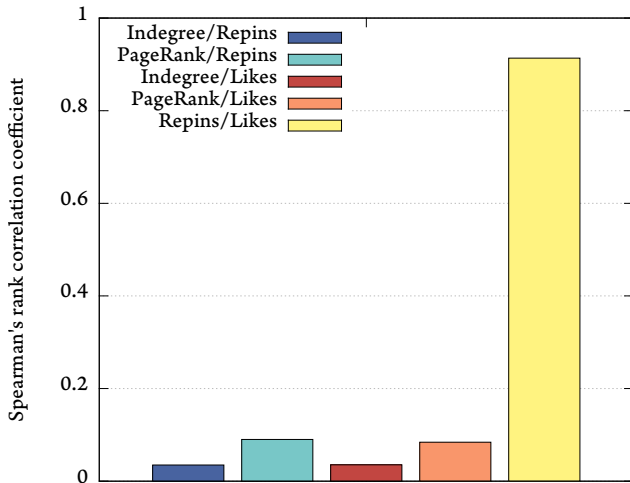


correlation of indegree with repins or likes is indeed even weaker for the top 10th percentile of users

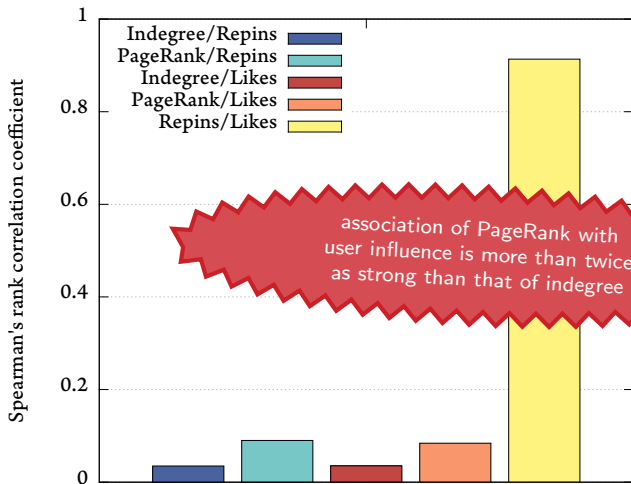
Rank correlation for the top 10th percentile of users



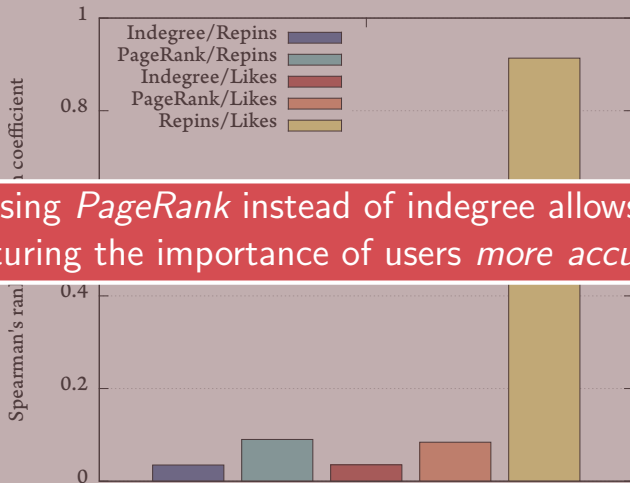
Rank correlation for the top 1st percentile of users



Rank correlation for the top 1st percentile of users



Rank correlation for the top 1st percentile of users



Using *PageRank* instead of indegree allows for capturing the importance of users *more accurately*.

References

- [CHBG10] Meeyoung Cha, Hamed Haddadi, Fabrício Benevenuto, and P. Krishna Gummadi, *Measuring User Influence in Twitter: The Million Follower Fallacy*, ICWSM, 2010.
- [LSMW98] Page Lawrence, Brin Sergey, Rajeev Motwani, and Terry Winograd, *The PageRank Citation Ranking: Bringing Order to the Web*, Technical report, Stanford University, 1998.
- [ZSS⁺14] Changtao Zhong, Mostafa Salehi, Sunil Shah, Marius Cobzarenco, Nishanth Sastry, and Meeyoung Cha, *Social Bootstrapping: How Pinterest and Last.fm Social Communities Benefit by Borrowing Links from Facebook*, WWW, 2014.
- [ZSS13] Changtao Zhong, Sunil Shah, Karthik Sundaravadivelan, and Nishanth Sastry, *Sharing the Loves: Understanding the How and Why of Online Content Curation*, ICWSM, 2013.

Conclusion - Future Work

- The study of *influence patterns* is essential for the design of successful *advertising strategies*.
- We performed an in-depth *analysis* of *Pinterest*.
- We found that there is very little *correlation* between the ranking of users based on their *indegree* and their ranking based on the number of either *repins* or *likes* they receive.
- We proposed the use of *PageRank* instead of *indegree* for the identification of influential users.
- We found that *Pagerank's* correlation with the ranking of users based on *repins* or *likes* received is *limited*, however, *much stronger* than that of *indegree*.

thank you!

for further details visit:

<http://hive.di.uoa.gr/network-analysis/>

or email me at: sioutis@cril.fr