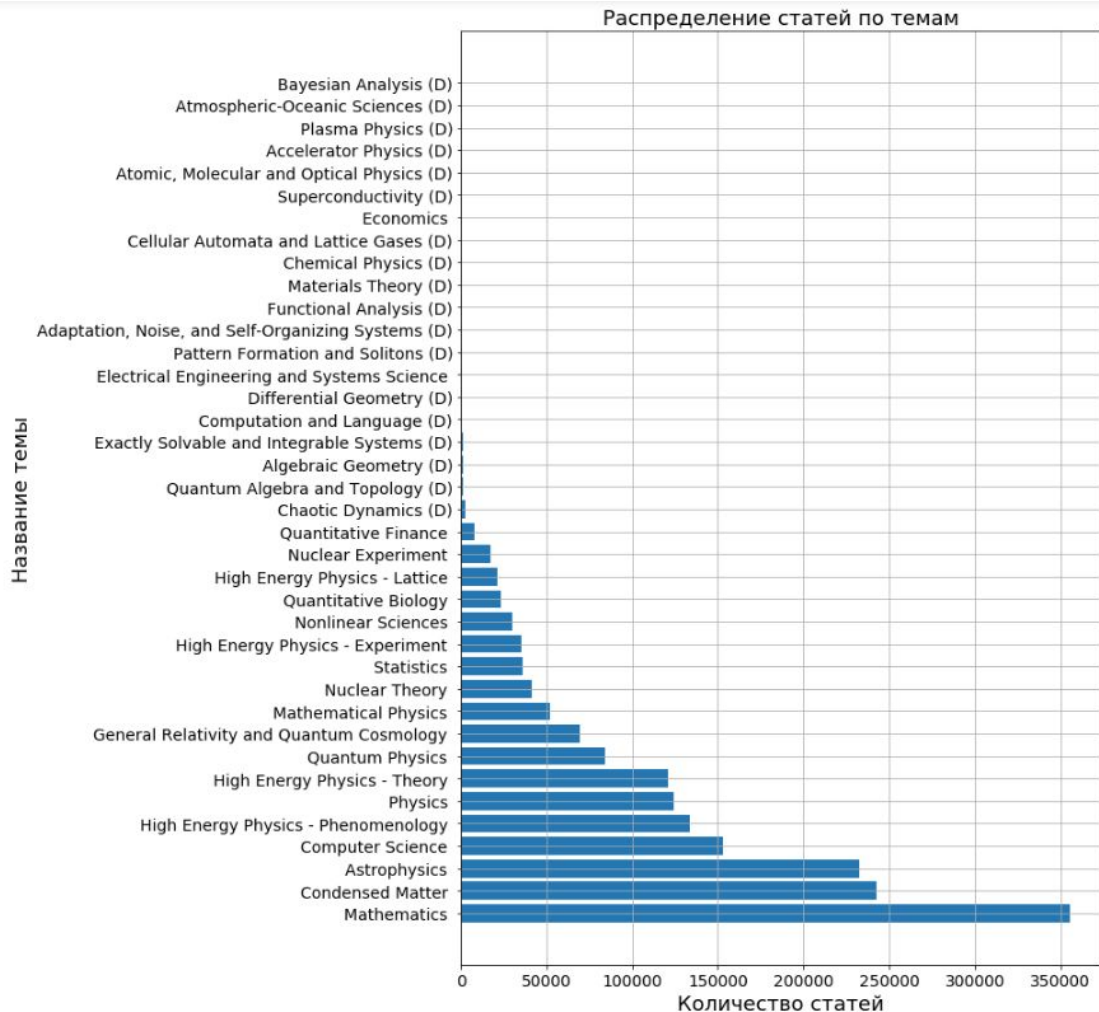


PZAD. Анализ графа Arxiv

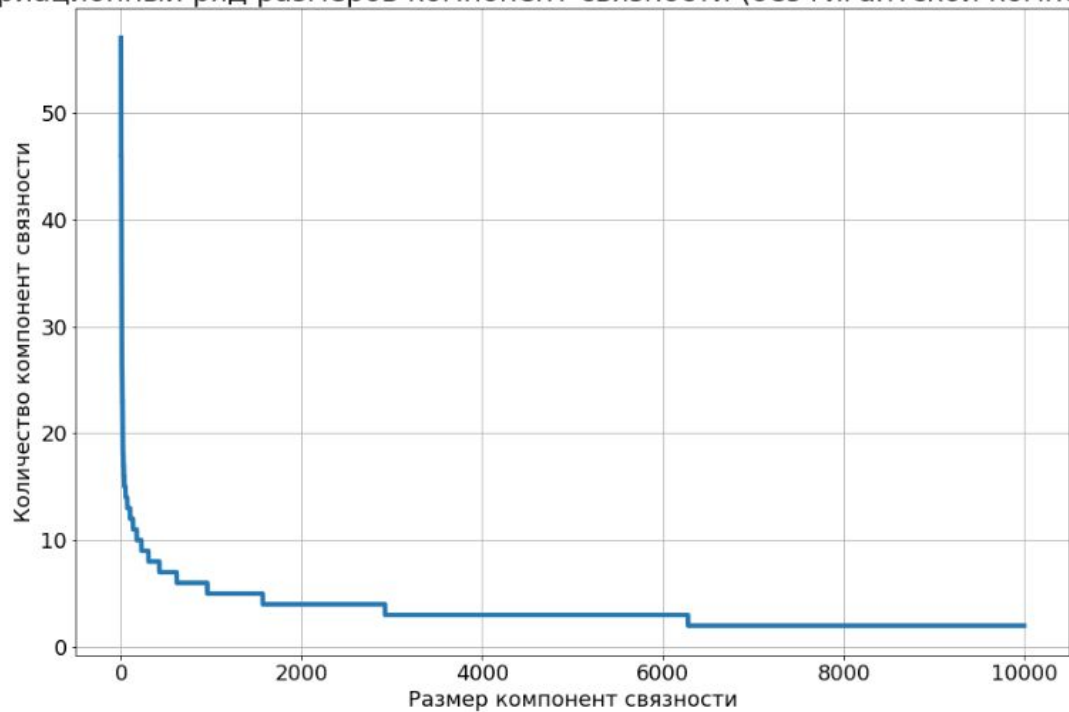
Урванов Егор

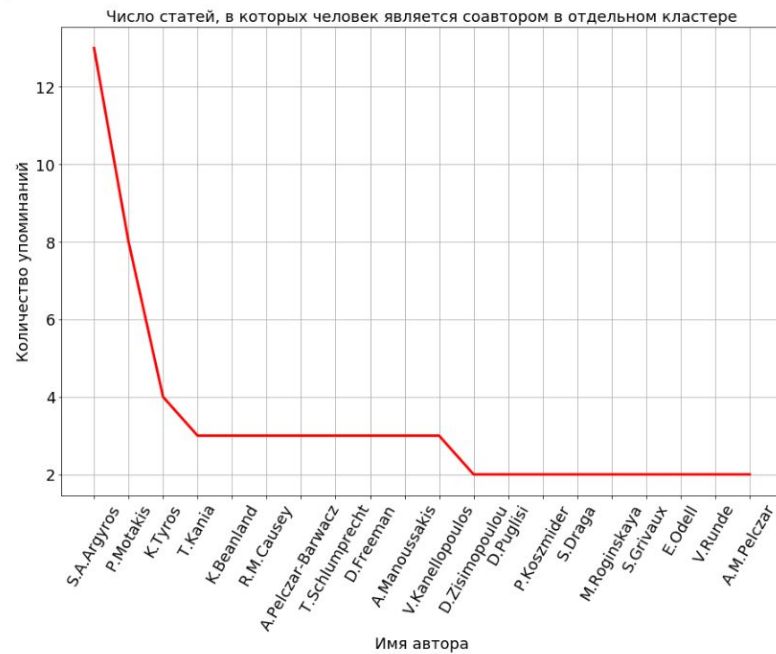
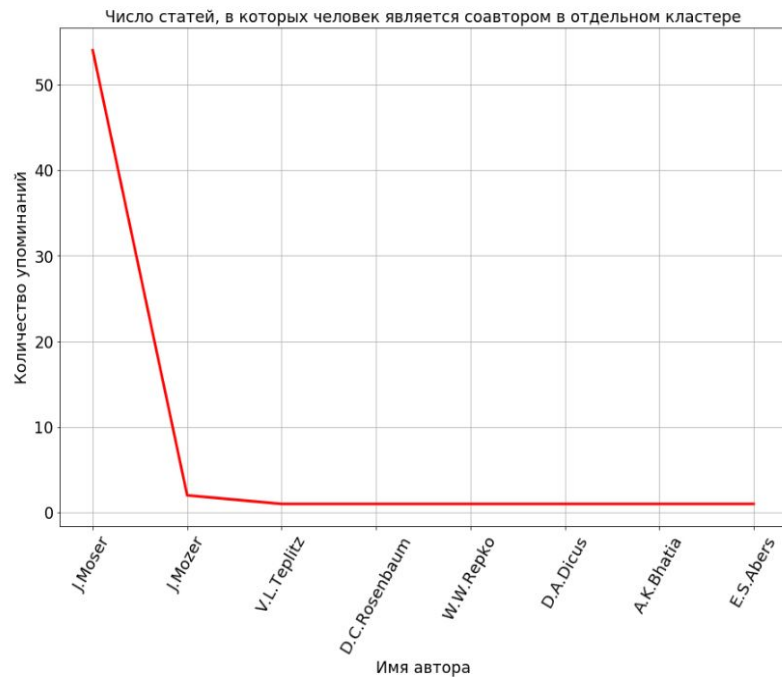
10 декабря 2018

- Устаревшие категории
- Новые категории



Вариационный ряд размеров компонент связности (без гигантской компоненты)



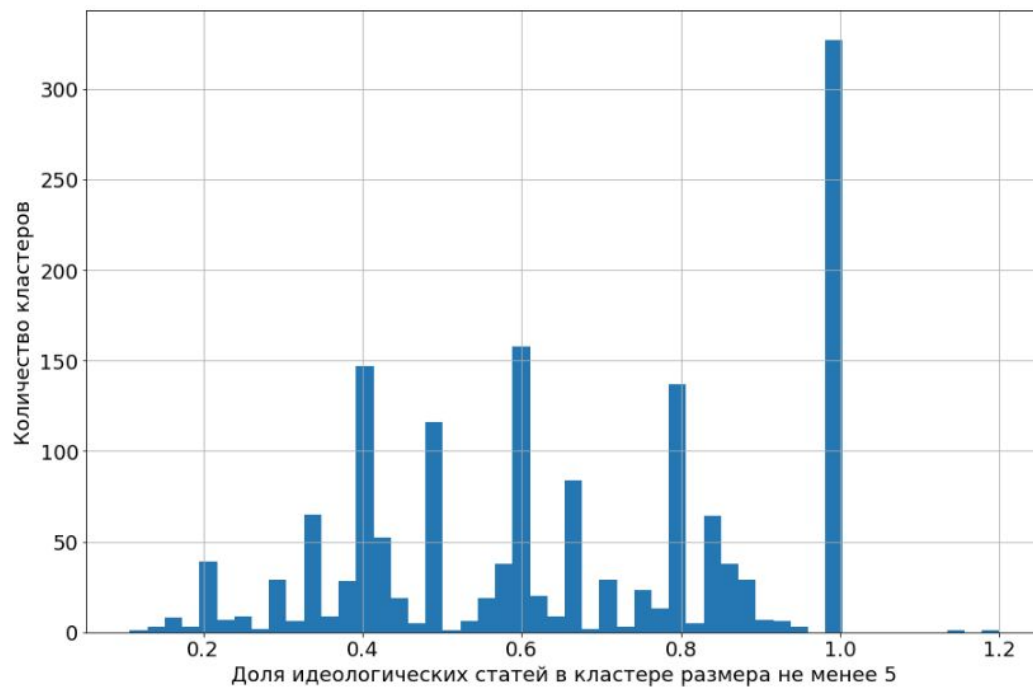


В кластере есть один или несколько идеологов

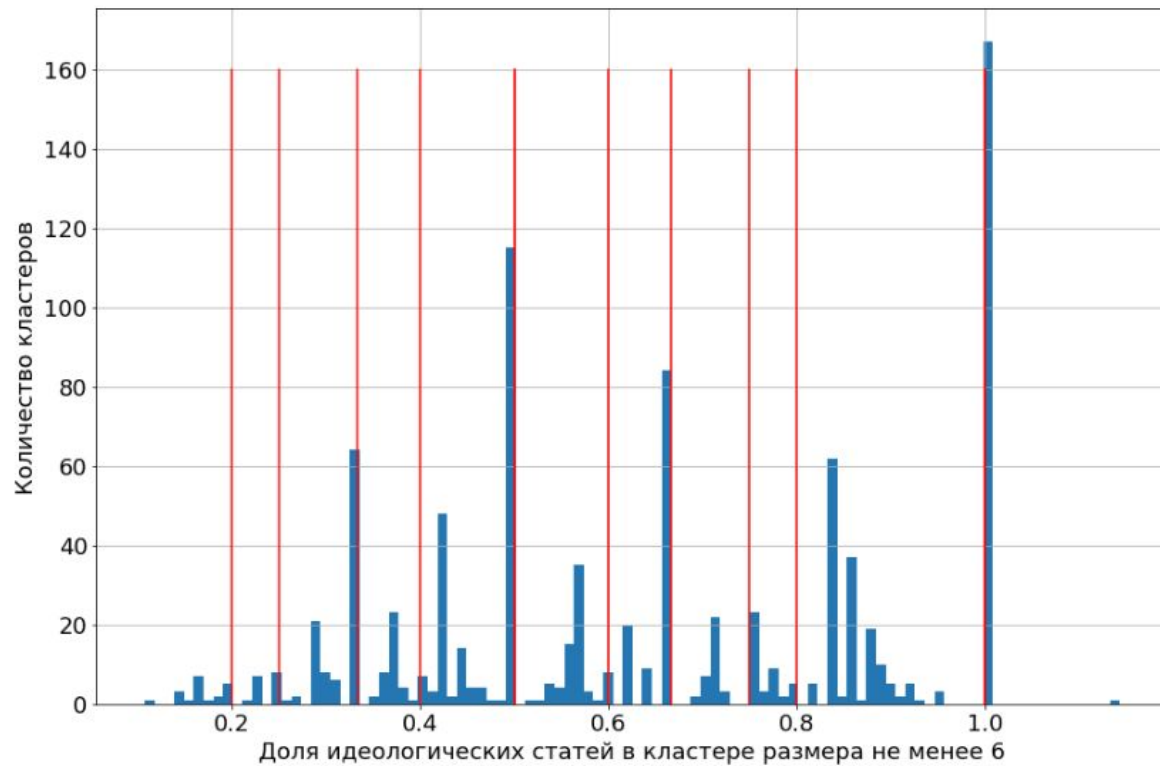
Такая тенденция  
наблюдается во многих  
кластерах



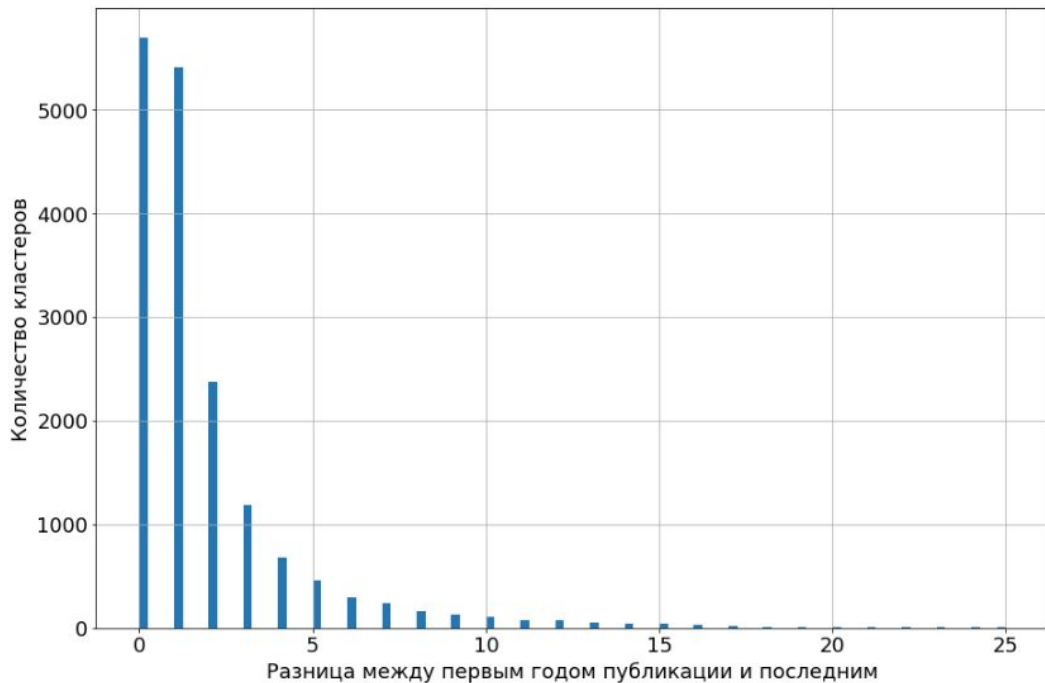
Видны спайки. Почему они имеют регулярную структуру?



Красные линии -- дроби  
вида  $n / m$ .



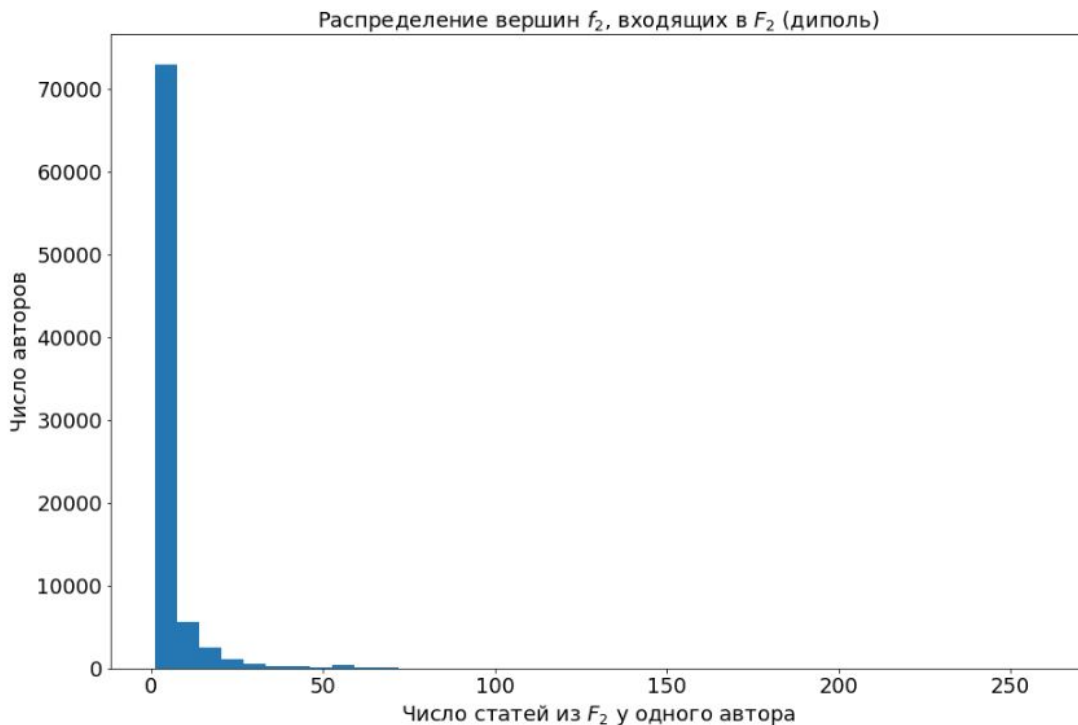
- \* Автор написали несколько статей когда-то в прошлом и больше не пишут статьи по данной теме
- \* Авторы написали несколько статей и ещё не успели написать других статей
- \* Автор написал много статей, но мало публиковал на arXiv (т. е. есть ссылки на внешние источники)
- \* Погрешность при составлении датасета



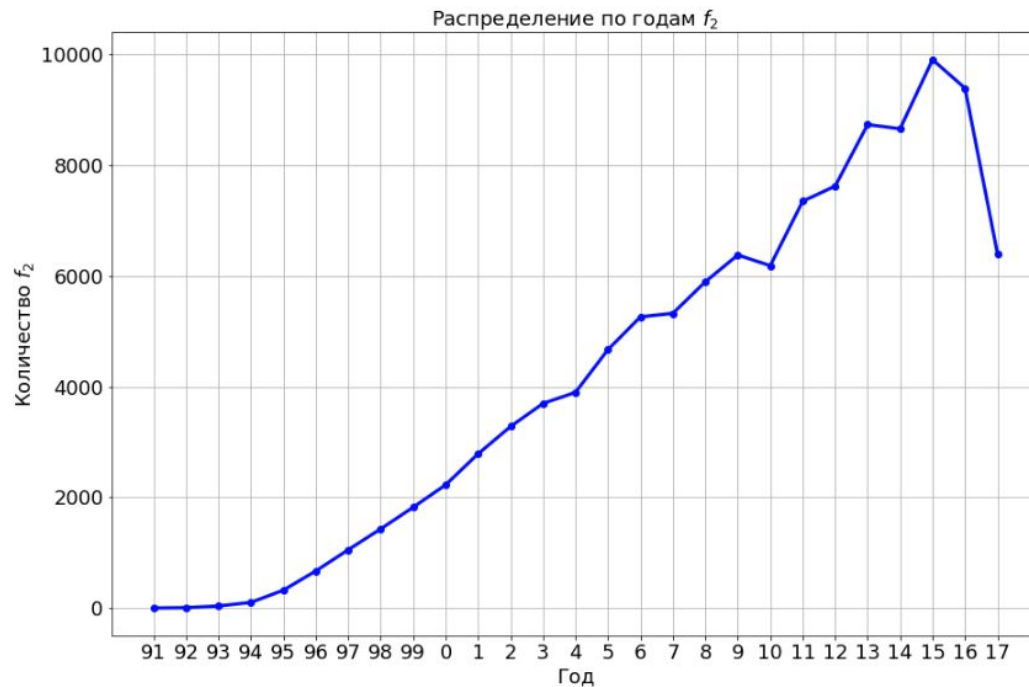


Число статей из диполей у одного и того же автора.

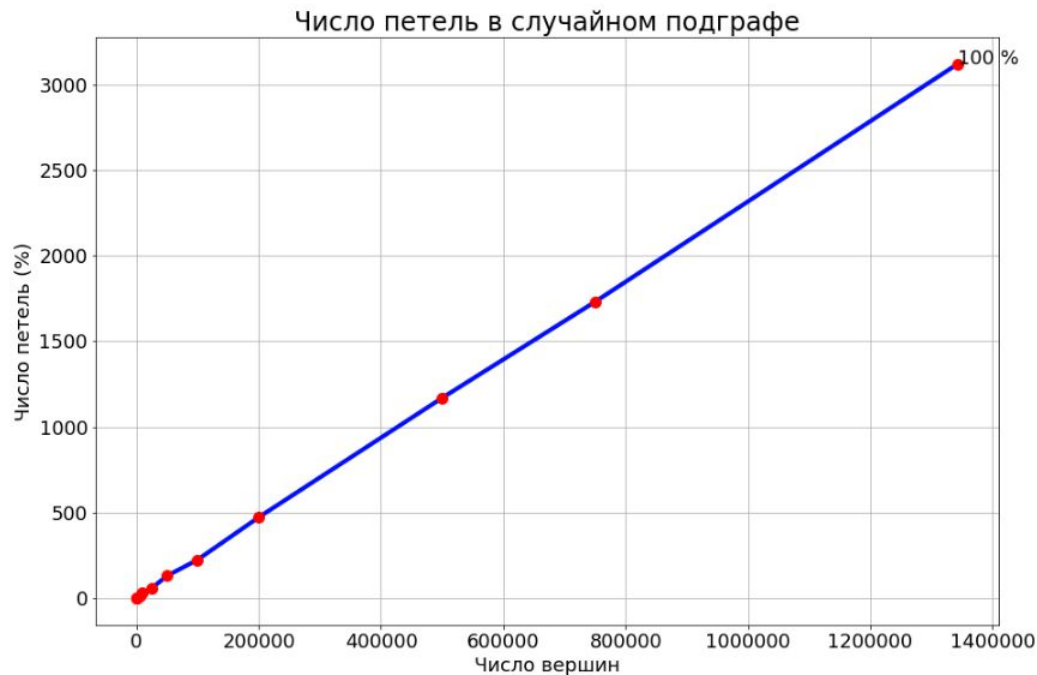
- У некоторых авторов более 150 статей
- Авторы, действительно, ссылаются друг на друга в “прошлом” и в “будущем” одновременно и это очень странно.
- Есть аномальная конференция “Planck”
- Число диполей растёт линейно, в зависимости от числа вершин (для стационарного графа)



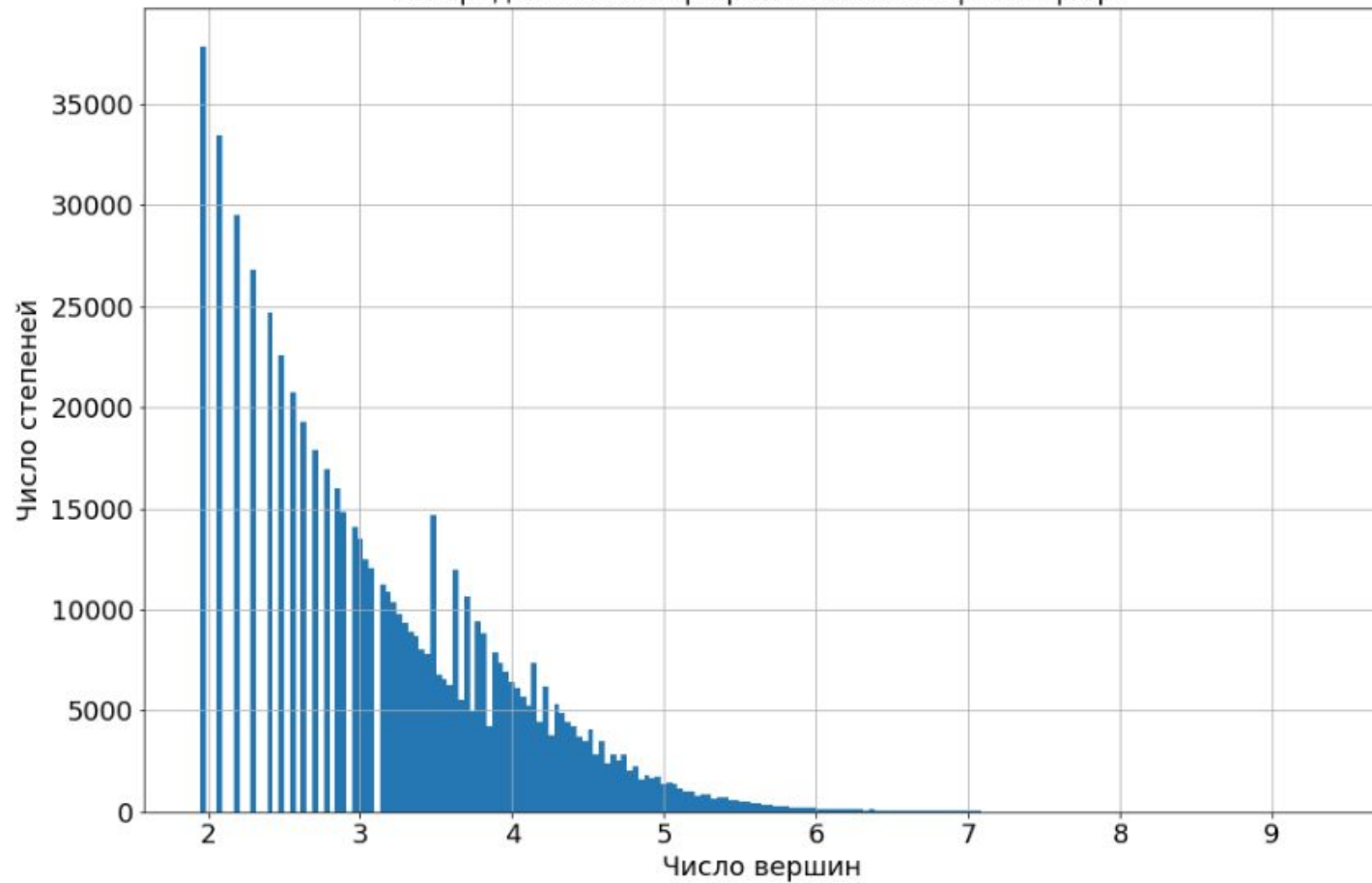
- Число диполей растёт со временем
- Начиная с 2004--2007 годов, распределение оказывается в большей степени случайным, чем ранее



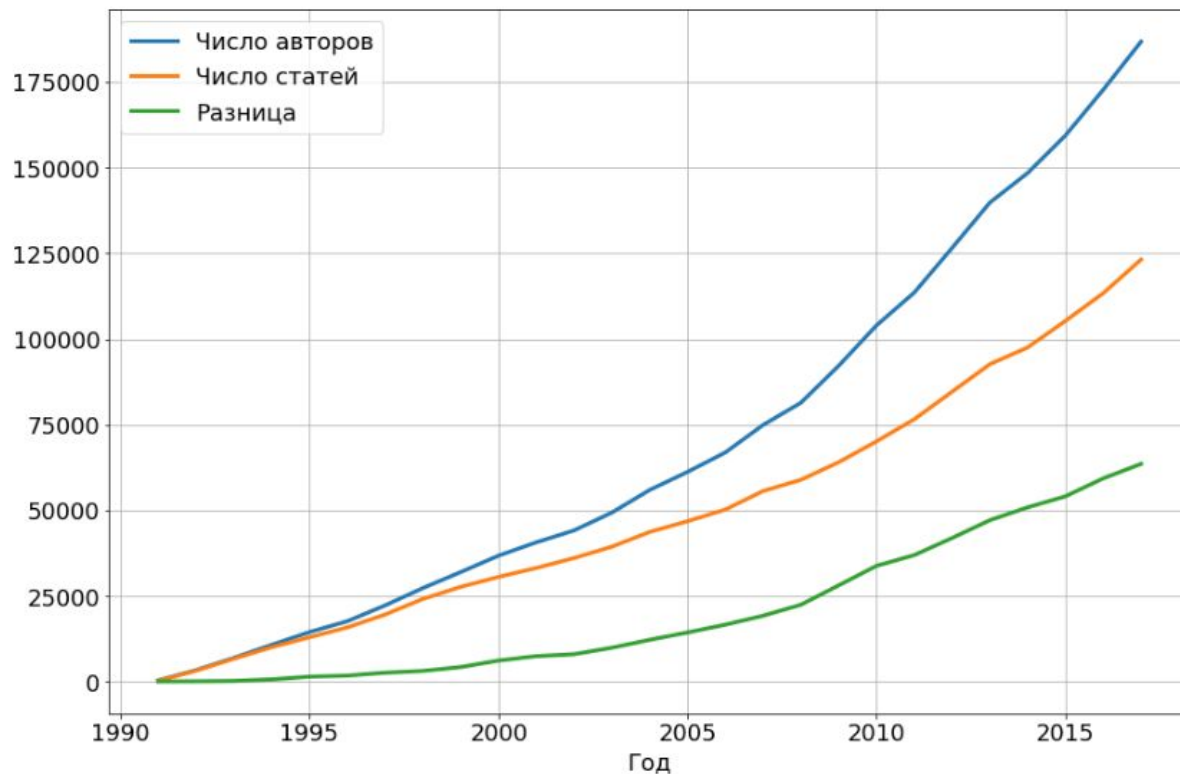
- Петли порождаются ошибками в графе
- Петли порождаются ссылками на другие версии статей
- Петли порождаются конференцией Planck

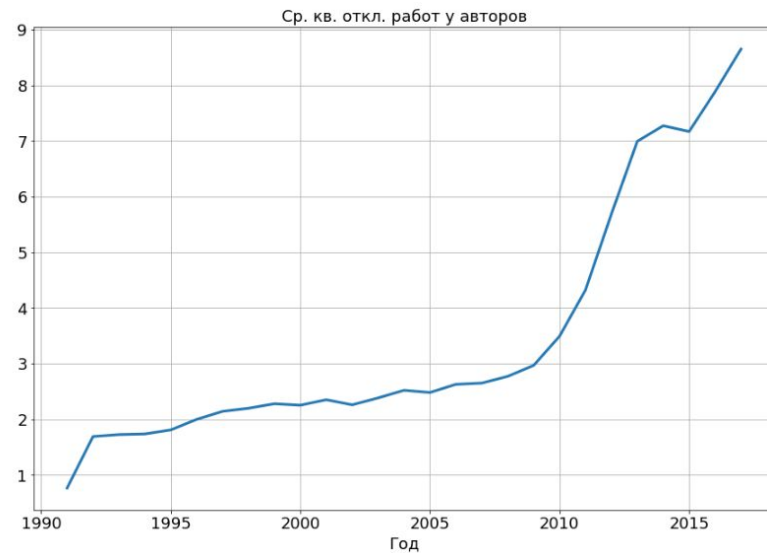
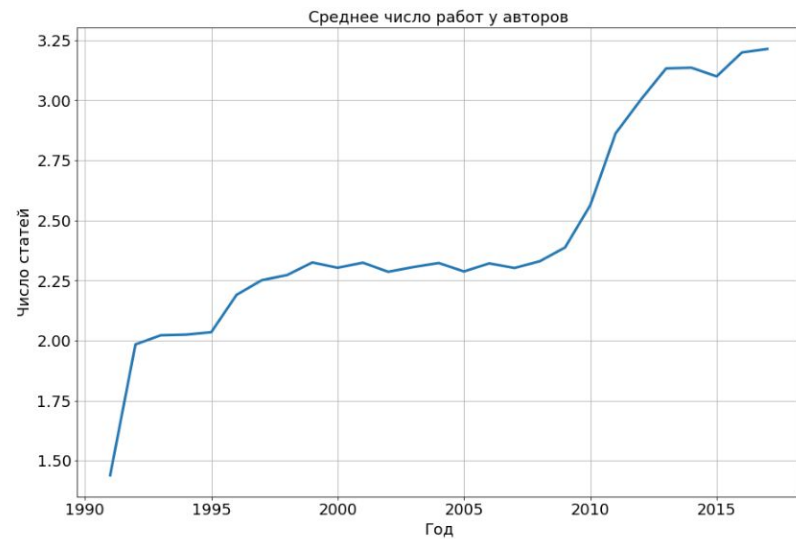


Распределение логарифма степеней вершин графа



- Число активных авторов растёт. Причём происходит это нелинейно
- Всё больше людей втягиваются в научные исследования (или просто начинают юзать arXiv)
- Становится всё больше нахлебников





Всплеск интереса к науке в 2009 - 2013 годах

Гипотеза малого мира  
выполняется

