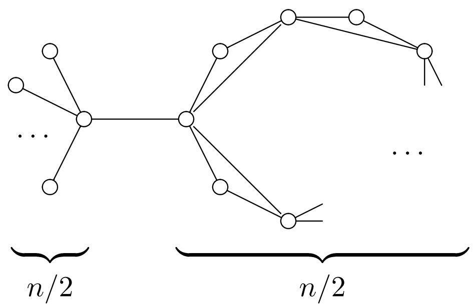
Домашнее задание 1

Задача 1. (2 балла). Приведите пример последовательности плотных (неразреженных) графов, степени вершин которых не подчиняются степенному закону. Приведите пример последовательности разреженных графов, степени вершин которых не подчиняются степенному закону.

Задача 2. (2 балла). Приведите пример последовательности неориентированных графов (кратные рёбра и петли разрешаются), степени вершин которых подчиняются какому-нибудь степенному закону. Обязательно доказательство степенного закона.

Задача 3. (1 балл). Приведите пример последовательности связных графов, которые не уязвимы к атакам на хабы.

Задача 4. (2 балла). Рассмотрим при растущих значениях n, кратных четырём, последовательность графов следующего вида:



Покажите, что при $n \to +\infty$ средний локальный кластерный коэффициент такого графа можно ограничить снизу положительной константой, а глобальный кластерный коэффициент стремится к 0.

Задача 5. (3 балла). Докажите, что при вероятности ребра p такой, что $pn^{\frac{2}{5}} \to 0$ при $n \to +\infty$ в случайном графе $G(n,p)$ в модели Эрдеша-Реньи асимптотически почти наверное нет клик на 6 вершинах, а при вероятности ребра такой, что $pn^{\frac{2}{5}} \to +\infty$ при $n \to +\infty$ в $G(n,p)$ асимптотически почти наверное есть клика на 6 вершинах.

Задача 6. (3 балла). Докажите, что при $p = \frac{c}{n}$, где $c < 1$, в случайном графе $G(n,p)$ в модели Эрдеша-Реньи асимптотически почти наверное каждая компонента связности содержит не более одного цикла.

Задача 7. (4 балла). В файле `graph.txt` находится фрагмент веб-графа 2002 года, заданный списком ориентированных рёбер, представляющих гиперссылки между страницами. В графе 875 тыс. вершин и 5.1 млн. рёбер. Постройте график кумулятивного[1] распределения полных степеней вершин этого графа в логарифмических координатах. Аппроксимируйте[2] «основную» часть графика степенным законом $\frac{c}{d^\gamma}$. Постройте на том же графике полученную функцию.

Задача 8. (4 балла). Коэффициент Жаккара для двух множеств A и B определяется как $J(A,B) = \frac{|A \cap B|}{|A \cup B|}$. Определим величину $J(e)$ для ребра $e = (u,v)$ графа G как $J(N_u, N_v)$, где за N_x обозначено множество соседей вершины x графа G. Далее, определим величину $J(G)$ как среднее значение $J(e)$ по всем ребрам графа G. При $n = 300$ сгенерируйте при разных значениях $p \in \{0, 0.05, 0.1, 0.15, \ldots, 1.0\}$ граф $G(n,p)$ в модели Эрдёша–Реньи и постройте график $J(G(n,p))$ в зависимости от p. Найдите теоретическую оценку для среднего значения этой величины и постройте её на том же графике.

Задача 9. (4 балла). Выберите LiveJournal какого-нибудь блогера. По адресу

`http://www.livejournal.com/misc/fdata.bml?user=<username>`

можно скачать список пользователей, которые являются друзьями или находятся в друзьях у пользователя `<username>`. Для каждого из друзей выбранного блогера также скачайте список его друзей, в итоге получится ориентированный граф друзей пользователя. Найдите локальный кластерный коэффициент выбранного блогера в графе социальной сети. Визуализируйте полученный фрагмент графа так, чтобы размер вершин и подписей к ним были пропорциональны полной степени. Вершины должны быть подписаны именами соответствующих пользователей.

[1] В случае кумулятивного распределения считается не число вершин заданной степени, а число вершин со степенью не меньше заданной.

[2] В `Python` это можно сделать, например, с помощью функции `curve_fit` модуля `scipy.optimize`.