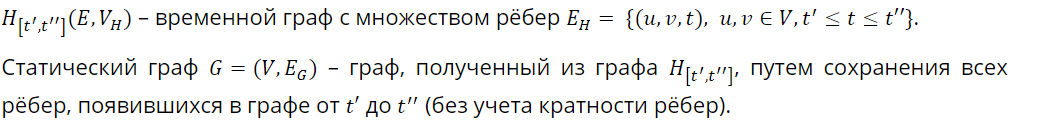
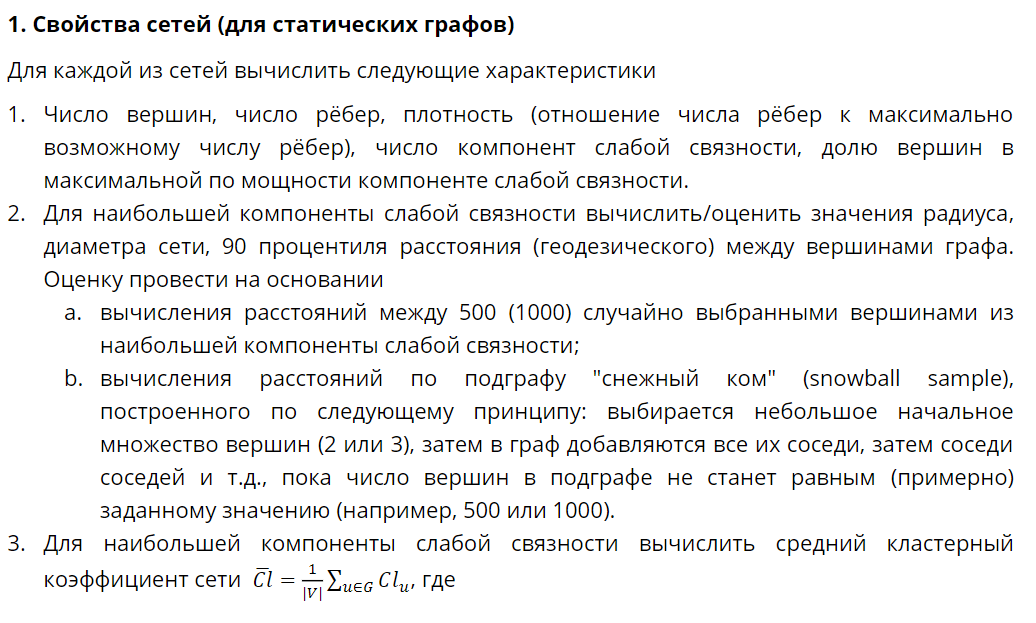
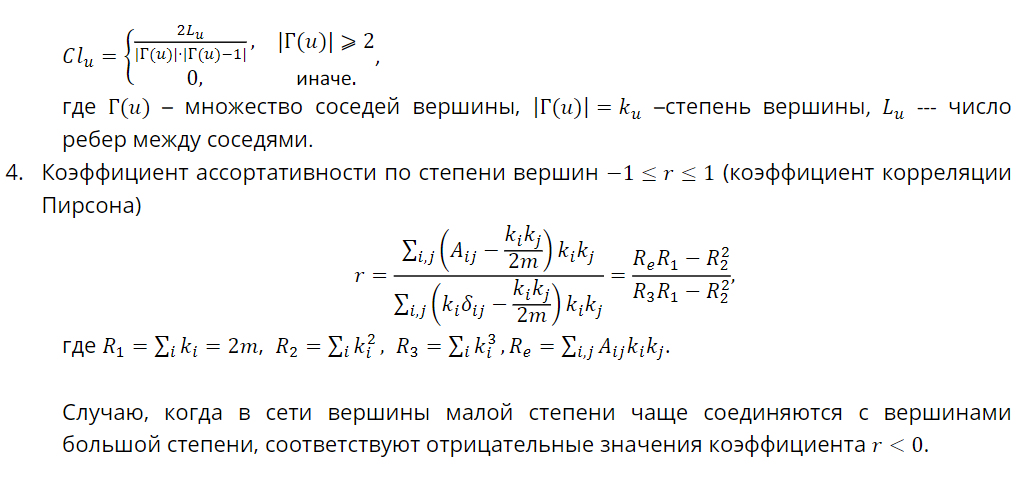
Задание основано на вычислении метрик рёбер и применении их для предсказания появления рёбер в темпоральных (временных) графах. ИдеязаданиявзятаизстатьиBruin, Gerrit Jan de; Veenman, Cor J.; van den Herik, H. Jaap; Takes, Frank W. (2021): Supervised temporal link prediction in large-scale real-world networks. In Soc. Netw. Anal. Min. 11 (1). DOI: 10.1007/s13278-021-00787-3.

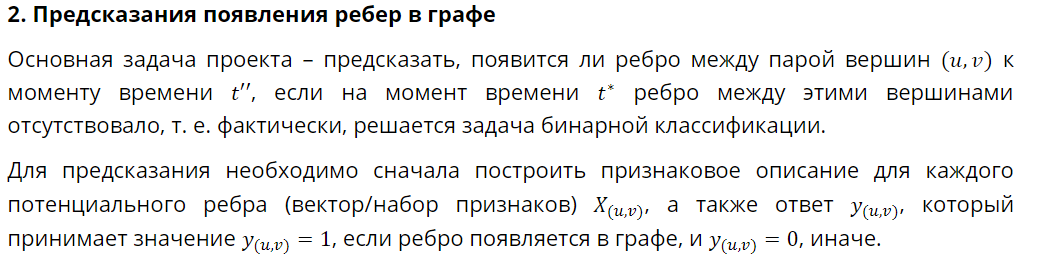
**Данные**

Для работы предлагаются датасеты, размещенные на <http://konect.cc/networks/> и представляющие собой реальный сети различной природы (социальные, информационные, технологические). Каждый датасет – это темпоральный (временной) граф, где для каждого ребра указана временная отметкав промежутке от t′ до t′′– время появления ребра в графе. Все графы считать неориентированными.









1. Построение векторов признаков для предсказания появления ребер в графе.

В работе [1] предлагается несколько вариантов построения признакового описания, а затем сравнивается качество полученных предсказаний. Нужно построить два различных набора признаков: статический (I) и один из II-A(см. Статью).

Б. Бинарная классификация.

Для обучения модели применяется алгоритм логистической регрессии. Для оценки качества построенной модели используется метрика AUC(Area Under the Receiver Operating Curve) – площадь под ROCAUC–кривой. ROСAUC–кривая строится на основании соотношения доли верно классифицированных объектов, обладающих некоторым свойством(TPRtruepositiverate) и долио бъектов, не обладающих свойством, но ошибочно классифицированных как обладающие этим свойством (FPRfalsepositiverate), при различных уровнях порога принятия решения.

TPR–доля рёбер, которые появились в графе y(u,v)=1 и которые классификатор отметил, как появившиеся в графе.

FPR–доля рёбер, которые не появились в графе, но которые были отмечены, как появившиеся.