

МЕТОДЫ ВЫДЕЛЕНИЯ СООБЩЕСТВ В СОЦИАЛЬНЫХ ГРАФАХ

Никишин Евгений Сергеевич

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: nikishin.evg@gmail.com

Социальным называют граф, обладающим следующим набором свойств:

- большинство (как правило, не менее 98%) вершин графа принадлежат одной компоненте связности
- диаметр этой компоненты невелик (около 5–6 рёбер)
- для вершин верно следующее утверждение: если А соединена с В и А соединена с С, то с высокой вероятностью также будут соединены В и С

Самым ярким примером социального графа является граф знакомств людей. В социальных графах вершинам свойственно объединяться в группы, называемые сообществами, внутри которых связей больше, чем вне. То есть граф представляет собой набор сообществ.

Предметом данной работы является исследование различных методов обнаружения таких сообществ. Рассматриваются как классические подходы [1], так и самые последние, учитывающие потенциально огромное количество вершин и рёбер в графах.

Разделяют алгоритмы, находящие непересекающиеся и пересекающиеся сообщества (т.е. запрещающие и допускающие принадлежность вершины сразу к нескольким группам) [2]. Всего в работе сравниваются 7 методов для обнаружения непересекающихся и 4 метода для обнаружения пересекающихся сообществ.

Численные эксперименты ставятся на искусственно сгенерированных данных и на реальных графах социальных сетей. Получены выводы о достоинствах и недостатках методов, а также о границах их применимости, предоставляя тем самым практическое руководство по выбору алгоритма обнаружения сообществ.

Литература

1. Fortunato S. Community detection in graphs // Physics reports, Volume 486, Issues 3–5, 2010, P. 75–174.

2. Xie J., Kelley S., Szymanski B. K. Overlapping Community Detection in Networks: The State-of-the-Art and Comparative Study // ACM Computing Surveys, Volume 45, Issue 4, 2013, Article № 43.