Аналитический обзор моделей случайных графов для моделирования социальных сетей

Сухов Дмитрий Юрьевич
студент 1го курса магистратуры
Московского технологического университета
г. Москва
E-mail: strongbad88@mail.ru

Понятие «Социальная сеть» впервые было использовано социологом Джеймсом Барнсом в 1954 году, задолго до появления привычных современному пользователю социальных сетей. Современное понятие этого термина можно описать как круг знакомых человеку лиц, где человек — центр социальной сети, его знакомые — ветви социальной сети, а отношения между людьми — связи.

В современном интернете под социальной сетью понимают площадку, на которой люди могут взаимодействовать друг с другом и объединяться в сообщества по общим интересам или по иным причинам для непосредственного общения между собой.

В настоящее время для построения различных моделей социальных сетей применяют теорию случайных графов. Существует много видов моделей, генерирующих случайные графы, близкие по свойствам к реальным сетям [1].

В своих работах [2], [3], [4] Барабаши и Альберт, а также Х. Джеонг дают описание статистикам интернета, которые послужили основой для создания науки о росте интернета. При ближайшем же рассмотрении можно заметить, что большинство реальных сетей, таких как биологические, транспортные и им подобные, имеют топологию, схожую с интернет сетью.

Для начала определимся с тем, что мы будем понимать под интернет сетью. Сеть интернет — так называемый интернет-граф, вершины которого представляют собой обособленную структурную единицу, расположенную в сети интернет. Интернет-граф формируют добавлением к нему новых вершин, соединяемых ребрами со старыми вершинами. Для большей наглядности возьмем за вершину интернет-графа сайт. В качестве ребер будут выступать ссылки, которые объединяют сайты, при этом ребра разумно считать направленными.

Таким образом, интернет-граф ориентирован, и он может иметь кратные ребра, петли и даже кратные петли (ссылки, расположенные на сайте, могут вести с одной страницы сайта на другую его страницу) [5]. Для подобной модели построения модель графов Эрдеша — Реньи является неактуальной.

Перечислим основные моменты исследования Барабаши — Альберт. Во-первых, интернет-граф — это весьма разреженный граф. У него на t вершинах примерно kt ребер, где $k \geq 1$ — некоторая константа. Для сравнения, у полного графа на t вершинах $C_t^2 = \theta(t^2)$ ребер. Во— вторых, — диаметр интернет-графа исключительно скромен.

К концу 20 — началу 21 века, интернет граф имел величину 5 — 7. [4] Это широко известное свойство социальной сети, которое обычно описывают выражением «мир тесен». Оно гласит о том, что любые два человека в мире разделены не более чем пятью уровнями общих знакомых, или же 6 уровнями связей. Аналогичная схема работает и для интернет сайтов: переходя по ссылкам, можно с любого сайта на любой другой перейти за 5 — 7 кликов компьютерной мыши. Но есть и некоторые ограничения. Так, к примеру, только-только созданные сайты могут не иметь

непосредственной связи с внешними ресурсами.

В-третьих, у интернет-графа весьма характерное распределение степеней вершин. Эмпирическая вероятность того, что вершина интернет-графа имеет степень d, оценивается как c/d^{λ}, где $\lambda \approx 2.1$, а c —нормирующий множитель, вычисляемый из условия «сумма вероятностей равна 1». Таким образом, интернет-граф имеет степенное распределение вершин, что, делает его очень похожим на многие реальные сети — все они подчиняются степенному закону, только у каждой из них свой степенной показатель распределения. Такие графы принято называть безмасштабными.

Критерий	Модель Эрдеша — Реньи	Модель Барабаши — Альберт
Разреженность графа	Плотный	Разреженный
Диаметр графа	Большой	Малый
Распределение степеней вершин графа	Нехарактерное	Характерное

Таблица 1 — сравнение моделей случайных графов

При изучении выбранных для сравнения критериев можно заключить, что модель Эрдеша — Реньи слабо применима для моделирования интернета и подобных сетей. Если подбором вероятности *p* можно добиться разреженности и тесноты (хотя и не с теми параметрами), то степенной закон совсем уж не имеет отношения к схеме Бернулли, в рамках которой появляются ребра обычного случайного графа. В модели *G (n,p)* степень каждой вершины случайного графа биномиальна с параметрами *n — 1* и *p*, и при тех p, которые хоть сколько-нибудь гарантируют разреженность (т.е. при $p = \theta(1/n)$), указанное биномиальное распределение аппроксимируется пуассоновским, а не степенным. Таким образом, можно заключить, что модель случайных графов Барабаши и Альберта является наиболее подходящей для моделирования социальных сетей.

Список источников

1. Райгородский *А.М.* Математические модели Интернета // Журнал «Квант». 2012. № 4.
2. Barabasi L.-A., Albert R. Emergence of scaling in random networks // Science. — 1999. — V. 286. P. 509–512.
3. Barabasi L.-A., Albert R., Jeong H. Scale-free characteristics of random networks: the topology of the world-wide web // Physica. — 2000. — V. A281. — P. 69–77.
4. Albert R., Jeong H., Barab ási L.A.Diameter of the world-wide web // Nature. — 1999. — V. 401. P. 130–131.
5. *Райгородский А.М.* Модели случайных графов и их применение // Тр. МФТИ. 2010. Т. 2, № 4. С. 130–140.