

## АНАЛИЗ КОМПЬЮТЕРНЫХ СОЦИАЛЬНЫХ СЕТЕЙ НА ОСНОВЕ ГРАФОВЫХ МОДЕЛЕЙ

Мустафин С.Н. студент группы ИС-22  
Есмагамбетов Т.У. магистр, ст.  
преподаватель кафедры  
«Информационно-вычислительных систем»  
Карагандинский экономический университет  
Казпотребсоюза

Увеличивающаяся популярность компьютерных социальных сетей подтверждается огромным числом пользователей Facebook, Twitter, ВКонтакте, Google+ и др. Растущая доступность к Internet с помощью различных устройств (компьютеры, смартфоны, планшеты), позволяет большинству пользователей круглосуточно находиться в сети и создавать цепи виртуальных взаимоотношений. Поскольку социальные сети становятся предпочтительными средствами для общения людей, ожидается, что их структура будет все более и более зеркально отражать реальное общество и отношения.

Существенная особенность социальных сетей - дружеские отношения между участниками. Главным образом они состоят в разрешении просматривать список друзей друг друга, вести беседу и отправлять различный контент: новости, фотографии, ссылки и т.д. Такие разрешения могут быть как взаимными, так и наоборот.

Математически любая социальная сеть представляется в виде графа  $G = (V, E)$ , где  $V$  - множество вершин графа;  $E$  - множество ребер графа [1]. В графе социальной сети вершинами являются участники, а ребра означают наличие отношений между ними. Анализ взаимоотношений в компьютерных социальных сетях представляет интерес по нескольким причинам. Во - первых, полное изучение структуры больших реальных социальных сообществ раньше было невозможным или как минимум очень дорогим. Во - вторых, в виртуальных сообществах данные четко определены структурными ограничениями самой социальной сети, по сравнению с реальными жизненными отношениями. Непосредственный анализ взаимосвязей в виртуальных социальных сетях имеет ряд препятствий. Во - первых, масштабируемость - проблема, с которой сталкиваются все, кто хочет изучить большую социальную сеть. Во - вторых, должны быть введены правильные социальные показатели, в целях выявления и оценки особенностей рассматриваемой сети. В 2010 некоторые авторы [2] подсчитали, что полный граф Facebook содержал 44 терабайт данных. Если бы даже эти данные были получены и хранились локально, создать и реализовать эффективные функции, которые бы анализировали такой граф или оценивали простейшие его метрики - непосильная задача в настоящее время.

Если говорить о графовых моделях социальных сетей, то для их анализа иногда используют коэффициент плотности - отношение числа ребер в анализируемом графе к числу ребер в полном графе с тем же числом вершин (полный граф - это граф, вершины которого соединены между собой) [3]. Также, такие величины, как число путей заданной длины (последовательность вершин, связанных между собой), минимальное число ребер, удаление которых разбивает граф на несколько частей, могут являться характеристиками виртуальных социальных сетей. Для того чтобы определить вес вершин графа (т.е. насколько большое влияние в рамках конкретной сети имеет её участник), применяют такое понятие, как центральность - мера близости к центру графа [4,5]. Важно понимать, что в данном случае имеется в виду не геометрическая центральность при визуализации графа. Есть разные способы определения центральности, поэтому существуют различные её меры. Центральность по степени - количество связей, проходящих через вершину: входящие связи показывают популярность участника, выходящие - его общительность. Центральность по близости является показателем того, насколько быстро распространяется информация в сети от одного участника к остальным, то есть насколько близок рассматриваемый участник ко всем

остальным участникам сети. Центральность по посредничеству - показывает важность участника сети при распространении информации. Центральность по собственному вектору - показатель, демонстрирующий зависимость между центральностью участника и центральностями его друзей. 60 Определению связанных групп (сообществ) отводится большое значение при анализе социальных сетей. Основной характеристикой сообщества является наличие большого числа связей между входящими в них участниками и существенно меньшим числом связей с остальными участниками [6].

Самый простой пример сообщества: каждый участник связан с каждым, и в данную группу не могут быть включены другие участники сети, поскольку они не имеют связей со всеми членами сообщества. Такие группы ещё называют кликами. В теории графов, клика - это максимально полный подграф данного графа. При анализе процессов распространения информации в графах, появляется другое определение сообщества: множество участников, в котором путь между двумя любыми участниками имеет не более одной промежуточной вершины. Поэтому в связанной группе информация от одного участника к другому передается с минимальными искажениями. Целью определения сообществ является поиск регионов сети, в которых происходит активное взаимодействие участников. По алгоритму эта задача похожа на задачу о разделении графов [7].

На основе поведения связей между вершинами производится разделение сети на плотные регионы. Ввиду того что, виртуальные социальные сети динамичны, появляются затруднения с выявлением сообществ. В некоторых случаях удастся интегрировать информационное содержимое сети в процесс определения сообществ. К методам определения сообществ ещё относят многомерное шкалирование или факторный анализ матрицы связей. Для того чтобы сделать вывод о влиянии конкретных участников сети на всю сеть нужно изучить методы выявления и описания свойств сетей, закон распространения этих свойств, создание методов, которые позволяют определять причины взаимоотношений участников по структуре сети. Эти процессы крайне необходимы при анализе больших современных социальных сетей.

#### **Список использованной литературы:**

1. Страхов М.Д. Социальные сети и их влияние на современное общество // Молодёжный научно - технический вестник. 2015
2. Gjoka, M., Kurant, M., Butts, C., Markopoulou, A.: Walking in facebook: a case study of unbiased sampling of OSNs. In: Proc. of the 29th conference on Information communications, pp.
3. Бартенев М.В., Вишняков И.Э. Использование графовых баз данных в целях оптимизации анализа биллинговой информации // Инженерный журнал: наука и инновации. 2015
4. Blondel, V., Gajardo, A., Heymans, M., Senellart, P., Van Dooren, P.: A measure of similarity between graph vertices: Applications to synonym extraction and web searching. Siam Review pp. 647-666 (2014) 47
5. Чураков А.Н. Анализ социальных сетей // Социологические исследования. 2001. № 1. С. 109-121.
6. Социальная сеть. URL: [http://ru.wikipedia.org/wiki/Социальная\\_сеть](http://ru.wikipedia.org/wiki/Социальная_сеть).
7. Charu C., Social network data analytics, 2016.