

ПОСТРОЕНИЕ СОЦИАЛЬНОГО ГРАФА ПОЛЬЗОВАТЕЛЯ СОЦИАЛЬНОЙ СЕТИ

Нечаева П. А.¹

(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Объектом разработки является программный комплекс для построения и визуализации социального графа пользователя социальной сети. Целью работы является реализация приложения для поиска и визуализации социального графа пользователя социальной сети с минимальной входной информацией о запрашиваемом пользователе. В работе описываются основные методы разработки программного комплекса. В ходе разработки программного комплекса были использованы элементы теории графов, методы визуализации социальных графов. Для визуализации графов был выбран силовой алгоритм компоновки графов. Кроме того, в докладе рассматриваются методы работы с социальными сетями ВКонтакте и Одноклассники. Данная работа основывается на модификации силового алгоритма визуализации графов. Алгоритм позволяет визуализировать эстетически привлекательный, интуитивно понятный и интерактивный граф. Результатом работы является построенный социальный граф пользователя социальной сети.

Ключевые слова: социальный граф, социальные сети, методы визуализации графов, силовой алгоритм визуализации графов.

1. Введение

Социальные сети в настоящее время пользуются большой популярностью среди всех поколений. В социальных сетях на данный момент находится большое количество информации о человеке. Помимо даты рождения, родного города, круга интересов информация из социальных сетей предоставляет доступ к друзьям пользователя. На основе знаний о друзьях человека можно провести исследование о его кругах общения. Данная информация представляет ценность как с точки зрения социальных наук, так и со точки зрения математических наук.

На данный момент интернет-пространство предоставляет доступ ко множеству социальных сетей. Некоторые социальные сети созданы для общения между пользователями. Другие носят

¹ Полина Александровна Нечаева, инженер-программист (polinaa12@gmail.com).

характер видео-хостингов или предназначены для обмена фотографиями и короткими видео. Однако подавляющее большинство популярных социальных сетей сочетают в себе обширный набор функций. Такие социальные сети и планируется изучить в данной работе.

Наиболее популярными социальными сетями в России, подходящими для данного исследования, являются ВКонтакте и Одноклассники [1]. Данные социальные сети обладают информацией о пользователе, фотографиями, позволяют получать доступ к кругу его общения.

Цель данной работы заключается в разработке программного комплекса для построения и визуализации социального графа пользователя социальной сети. Основными подпрограммами программного комплекса выступают: подпрограмма поиска пользователя в социальных сетях ВКонтакте и Одноклассники; подпрограмма сбора данных о выбранном пользователе в определенной социальной сети, а также подпрограмма для построения социального графа пользователя.

В данной работе рассматриваются методы визуализации социального графа. Критерием выбора определенного алгоритма являются эстетическая привлекательность графа, читаемость графа, сложность алгоритма.

Для данной работы был выбран силовой алгоритм визуализации графа. На основе рассмотренного алгоритма реализуется программная часть работы.

2. Основная часть

2.1. ЭЛЕМЕНТЫ ТЕОРИИ ГРАФОВ

В данной работе рассматривается поведение пользователя социальной сети путем анализа построенного его социального графа. Для построения графа необходимо его определить.

Пусть существует непустое множество X , состоящее из n элементов. Зададим отображение $G - X$ множества на X . Пары элементов множества x_i и x_j будут соединены при условии, если x_j , принадлежащее G , связано с x_i , принадлежащим G , и наоборот.

Пусть V – множество отношений. Следовательно, (X, V) – это пара множеств, которая определяет сам граф [2].

Если количество вершин графа конечно, то граф называется конечным, иначе бесконечным. В данном исследовании рассматриваются только конечные графы.

Смежные вершины графа – это те вершины, которые соединены ребром. В данной работе «центральная» вершина будет смежной со всеми остальными вершинами множества X .

Основная задача исследования – определить наличие иных смежных вершин графа $G(V, X)$.

В визуализируемом графе определяются также и висячие вершины. Это те вершины, которые имеют одно инцидентное ей ребро. Условно говоря, друг заданного пользователя, не имеющих с ним общих связей (друзей).

2.2. ГРАФЫ В СОЦИАЛЬНЫХ СЕТЯХ

Социальную сеть можно интерпретировать как множество вершин, которыми являются социальные субъекты. Это могут быть пользователи, социальные группы, а также связи между ними. Субъекты социальной сети группируются по принципу сходства связей. Например, общие интересы, общие социальные группы или общий родной город. Принципы сходства связей могут быть любыми.

На основе данных из социальных сетей можно построить социальный граф пользователя или социальный граф интересов. Данные графы схожи по принципам, но имеют некоторые принципиальные различия. В социальном графе связи формируются на основе друзей пользователя, тогда как в графе интересов связи формируются по сходству увлечений. Связи в таком графе могут быть как «пользователь – пользователь» или «пользователь – интерес», так и «интерес – интерес».

В рамках данного исследования разрабатывается программный комплекс для построения социального графа пользователя социальной сети.

В ходе сбора данных из социальных сетей возникают две наиболее сложные проблемы. Во-первых, социальные сети не имеют единой топологии. Поэтому алгоритм сбора данных для

каждой сети необходимо подстраивать под их особенности. Во-вторых, встает проблема закрытости аккаунтов пользователей. Из-за этого некоторые социальные связи становятся неполными, так как отсутствуют полные данные о пользователе.

2.3. МЕТОДЫ ВИЗУАЛИЗАЦИИ ГРАФА

В данном исследовании происходит работа с динамическими графами. Для подобных графов наиболее популярными алгоритмами визуализации являются силовой и пружинный.

В работе эмпирическим путем был выбран силовой алгоритм визуализации графов. Эстетическая привлекательность графа обусловлена равномерным распределением длины ребер между собой, связанные вершины находятся вблизи друг друга, а несвязные, соответственно, отдалены друг от друга. Кроме того, силовой алгоритм позволяет прорисовывать промежуточные этапы построения графа, что дает возможность проследить за динамикой графа.

Однако силовой алгоритм имеет ряд недостатков, наиболее важный из которых – большое время выполнения. Силовые алгоритмы имеют итоговую сложность $O(n^2)$, где n – количество вершин графа. Существуют различные модификации алгоритма, которые позволяют сократить время выполнения до $O(n \log n)$, где n – количество вершин. Также одним из недостатков силового алгоритма является динамичность системы. Из-за действия сил притяжения на вершины со множеством связей система долгое время не может прийти в состояние равновесия, из-за чего происходит сильная подвижность подобных узлов.

Силовой алгоритм является интуитивно понятным ввиду теоретической основы метода. Как упоминалось ранее, силовой алгоритм визуализации моделирует физическую систему тел с силами, действующими между заданными телами [3].

При рассмотрении графа как физической системы тел вершины графа соответствуют заряженным частицам, а ребра графа аналогичны пружинам. На каждую вершину в системе действуют кулоновские силы и силы закона Гука. В том случае, если частицы заряжены одноименно и пружины растянуты, на них действуют силы притяжения согласно силам упругости

и силы отталкивания согласно электростатическим силам. Благодаря этим силам вершины динамического графа не склеиваются из-за достаточного отталкивания вершин и действия пружинных сил оставляют вершины в одном кластере. Однако при большом количестве пружин система долгое время не может прийти в состояние покоя, из-за чего происходит сильная подвижность узлов с множеством связей.

При достижении минимального энергетического состояния системы длина пружин будет одинакова, так как силы, действующие на вершины, компенсируют друг друга.

2.4. РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА

Для построения и визуализации социального графа пользователя социальной сети необходимо было разработать программный комплекс.

В качестве способа реализации программного комплекса было выбрано веб-приложение. Веб-приложение разработано на основе открытого веб-сервера Apache HTTP веб-сервер. Основными языками программирования являются Python и JavaScript.

Для разработки данного вида графа была использована библиотека языка JavaScript – VivaGraph. В исходном коде библиотеки были изменены следующие параметры: длина пружины, коэффициент закона Гука для притяжения узлов, коэффициент закона Кулона для отталкивания узлов, тета-коэффициент из моделирования Barnes Hut для оптимизации построения связей между вершинами, коэффициент силы сопротивления для установки жёсткости системы [4]. Значения параметров приведены в таблице 1.

Таблица 1. Настраиваемые параметры алгоритма

Параметры	Значение
Длина пружины	35
Коэффициент упругости	0,00055
Коэффициент сцепления	0,09
Коэффициент отталкивания	-1
Тета-коэффициент	0,1

Каждый параметр имеет значение. Длина пружины – идеальная длина для пружины в физической модели. Коэффициент упругости – закон Гука: чем выше коэффициент, тем меньше расстояние между вершинами и, соответственно, выше частота колебаний. Коэффициент сцепления, или коэффициент силы сопротивления, используется для замедления системы. Коэффициент отталкивания, или коэффициент Кулона, используется для отталкивания узлов. Следовательно, чтобы вершины не притягивались, необходимо установить отрицательное значение. Тета-коэффициент из моделирования Barnes Nut изменяется в диапазоне от (0, 1); чем ближе к 1, тем больше узлов придется пройти алгоритму.

Для сбора данных из социальных сетей был реализован доступ к API социальных сетей ВКонтакте и Одноклассники на языке Python. Однако API социальной сети Одноклассники не предоставляет полный доступ к необходимому функционалу для поиска пользователей по параметрам. Был изучен трафик посылаемых запросов к социальной сети, что позволило получить доступ к внутреннему API Одноклассников и осуществить поиск пользователя.

Основным функционалом приложения являются поиск человека по заданным параметрам в социальных сетях ВКонтакте и Одноклассники; сбор данных о выбранном пользователе, построение социального графа пользователя социальной сети.

Результатом выполнения программы является построенный граф (рис. 1). Граф, изображенный на рис. 1, построен на основе данных пользователя социальной сети ВКонтакте. Входные данные в формировании запросов к социальным сетям: имя, фамилия, пол, возраст, возрастной диапазон, страна, город, место работы. Обязательными полями являются только имя и фамилия.

Центральная вершина графа – искомый пользователь. Кроме того, вершины разделены по цветам. Наиболее темные вершины – мужчины, более светлые – женщины. Атрибутами вершины графа являются имена «друзей» пользователя. Визуально граф разделен на несколько кластеров по принадлежности к различным кругам общения. Также пользователь может производить

захват вершины графа и наблюдать за возвращением системы в состояние равновесия.

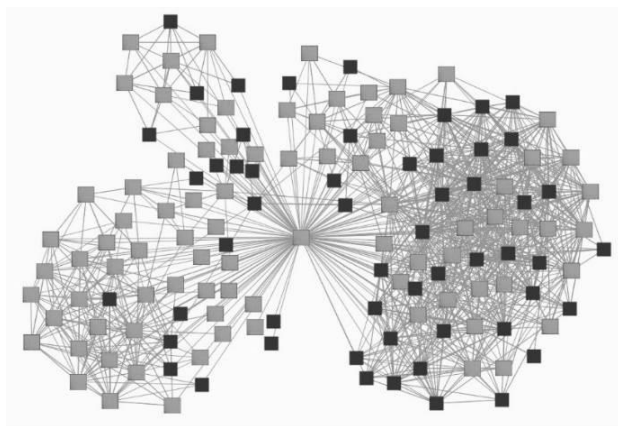


Рис. 1. Социальный граф пользователя социальной сети ВКонтакте

3. Заключение

Таким образом, был разработан программный комплекс для построения и визуализации социального графа пользователя социальной сети на основе модифицированного силового алгоритма.

Программный комплекс разработан в виде веб-приложения, результатом выполнения которого является социальный граф пользователя социальной сети в векторном формате. В качестве результата данного исследования выступает социальный граф пользователя социальной сети ВКонтакте и Одноклассники. Данный граф наглядно показывает, как распределяются по группам друзья заданного пользователя.

В ходе разработки программного комплекса был определен функционал разрабатываемого комплекса. В финале работы был получен полностью функциональный программный комплекс для построения социального графа пользователя социальной сети.

Литература

1. Аудитория интернета в России. [Электронный ресурс]. – URL: <https://webindex.mediascope.net/> (дата обращения: 11.02.2021).
2. БУРКОВ В.Н., ГОРГИДЗЕ И.А., ЛОВЕЦКИЙ С.Е. *Прикладные задачи теории графов. Т. 1.* – Тбилиси: Мецниереба, 1974.
3. MUELLER C., GREGOR D. P., LUMSDAINE A. *Distributed Force-Directed Graph Layout and Visualization // EGPGV.* – 2006. – Vol. 6. – P. 83–90.
4. TARAWANEH R.M., KELLER P., EBERT A. *A general introduction to graph visualization techniques // Visualization of Large and Unstructured Data Sets: Applications in Geospatial Planning, Modeling and Engineering – Proceedings of IRTG 1131 Workshop 2011.* – Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2012.

BUILDING A SOCIAL SOCIAL NETWORK USER GRAPH

Polina Nechaeva, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, software engineer (polinaa12@gmail.com).

Abstract: In this paper, the object of development is a software package for building and visualizing the social graph of a social network user. The aim of the work is to implement an application for searching and visualizing the social graph of a social network user with minimal input information about the requested user. The paper describes the main methods of developing a software package and the results of an experimental study of social connections of users of social networks. During the development of the software package, elements of graph theory, methods of visualizing social graphs were used. To visualize the graphs, a forceful graph layout algorithm was chosen. The report discusses methods of working with social networks VKontakte and Odnoklassniki. This paper is based on a modification of the force graph visualization algorithm. The algorithm allows you to visualize an aesthetically pleasing, intuitive and interactive graph. The result of the work is the constructed social graph of the user of the social network.

Keywords: social graph, social networks, graph visualization methods, force graph visualization algorithm

УДК 004.021 + 519.1

ББК 16.222